CERN-TH.5466/89

QCD Generators for LEP†

B. Bambah[1], J. Chrin[2], W. de Boer[3], J. Fuster[4],
J.W. Gary[5], M. Hahn[6], V. Khoze[7], W. Kittel[8], B. Lampe[4],
P. Mättig[4], T. Sjöstrand*[4], H. Stone[9], B. van Eijk[4]

*convenor
[1]L3/FBLJA Project, World Lab, CERN, Geneva, and
Dept. of Physics, Panjab University, Chandigarh, India
[2]Dept. of Physics, University of Manchester, United Kingdom
[3]Max-Planck-Institut, München, Federal Republic of Germany
[4]CERN, Geneva
[5]Physikalisches Institut, Universität Heidelberg, Federal Republic of Germany
[6]Institut für Exp. Kernphysik, Universität Karlsruhe, Federal Republic of Germany
[7]CERN, Geneva, and Leningrad Nuclear Physics Institute, Gatchina, USSR
[8]CERN, Geneva, and University of Nijmegen, The Netherlands
[9]University of Geneva, Switzerland

Abstract:
The current report contains an overview of QCD-related software that could be of interest for LEP physics studies. The main emphasis is on programs for the simulation of matrix elements, parton showers, fragmentation and decays. A number of Monte Carlo event generators are presented, mainly with respect to their physics contents. Overviews are given on the main concepts used in generators, and on comparisons between models and data. A standard set of common blocks for information interchange between event generators is presented.

---

†to appear in the Proceedings of the 1989 LEP Physics Workshop

CERN-TH.5466/89
July 1989

Contents

1 Introduction

The Z^0 dominantly decays into multihadronic final states. The experimental objectives will be to use multihadronic events to extract information on the electroweak (QFD) theory and on the strong (QCD) one, and to search for new particles and interactions. The complex structure of multihadronic events makes the detailed analysis more complicated than in purely leptonic/photonic final states. Good and reliable models are therefore needed to describe the transformation of primary quarks into the observable hadrons.

QFD is a theory which is solvable, at least in principle, through order-by-order perturbative calculations, given a set of input parameters. Not so for QCD. The fragmentation process (i.e. the process which transforms an initial set of partons into a final set of hadrons) has not yet been understood from first principles, but only in terms of vaguely QCD-inspired models, with many issues unsolved. Even in the large momentum transfer regime, where perturbative calculations can be used to describe jet production, the strong coupling constant α_S is large enough that yet uncalculated higher order corrections could well shift current results by a factor of two.

There will therefore be no lack of QCD-related subjects to be studied at LEP. The list includes:

- exploring the three-gluon vertex, checking the running of α_S, and performing other fundamental tests of QCD;
- studying multijet production (as a probe of perturbative QCD);
- looking for QCD coherence effects;
- gaining a better understanding of the fragmentation process; and
- finding signals for new physics or, at least, improving existing limits.

While it is important not to underestimate the opportunity for QCD studies at LEP, one should also remember that QCD is a subtle theory with few simple answers. One example: to determine the value of α_S to high precision would certainly be nice, but such a quantitative answer does not teach us anything qualitative about QCD, or about anything else. In this sense, the electroweak theory is much easier to explore, since e.g. Z^0 mass and width determinations have direct implications for the top and Higgs masses, and the number of neutrino species.

Several event generators have been developed to cope with QCD physics in e^+e^- annihilation. Due to the less well understood nature of the underlying theory, and to the larger variety of physical processes at play, the differences between the predictions of these generators is considerably larger than is the case for electroweak event generators [1]. That is fine; we do not strive to have every generator agree on the 1% level, since any such agreement would be entirely artificial and in no way reflect the underlying uncertainties.

Cross-sections vary rapidly, as a function of CM energy, around the Z^0. However, once a primary $q\bar{q}$ pair has been produced, the evolution into a hadronic final state follows the same basic rules, independent of CM energy. As a consequence, we can directly harness the experience which has been accumulated in the course of the PETRA/PEP/TRISTAN programs. This includes detailed experimental studies based on several of the generators described in this section. Of course, these programs evolve with time, so some studies will have to be repeated using the latest versions of the codes.

The multifaceted nature of multihadronic event generators makes writing a review of this kind rather challenging. Each of the major generators comprises a world of its own. Similarities exist, but differences abound. This will become apparent in reading the report. The similarities will be emphasized in section 2, where a general overview is given. The event generators themselves are described in sections 3 and 4.

In section 3 we describe the event generators which produce events with a limited number of partons according to given matrix elements, i.e. programs specialized to a very specific part of the full event generation. The main reason for this special attention is the rather confusing situation that has existed in the domain of full second order matrix elements. The main section of this report is section 4. Here a description is given of all multipartonic and/or multihadronic event generation programs we have found, which were not already discussed in section 3. This includes some programs which cover the full range of electroweak and perturbative QCD physics, fragmentation and decay, but also programs which specialize in only one or in a few aspects.

Past experience in the comparisons of event generators with experimental data is summarized in section 5. Here we also cover some of the main unresolved issues for the future. Work on trying to standardize the user interface to generators is covered in section 6. The final summary, in section 7, contains some value judgements of the programs on the market, as requested in the instructions for this Workshop.

Apologies to anyone who has been overlooked; one way of selling a program to the world, even outside Workshop times, is to submit it to the CERN program library. Obviously, all the programs described in this report either already are or soon will be included in the library – references to private files, where present, should hopefully be obsolete by the time this is read.

Finally, we remind the reader that a companion volume on 'Physics at LEP' is available. In particular, the two sections on QCD Tests [2] and Heavy Flavours [3] contain a wealth of information not covered in this report. In specific cases, references are given to these reports, but a full cross-reference at every point is clearly not possible.

References

[1] Report of the Electroweak Event Generators working group

[2] Report of the QCD working group

[3] Report of the Heavy Flavour working group

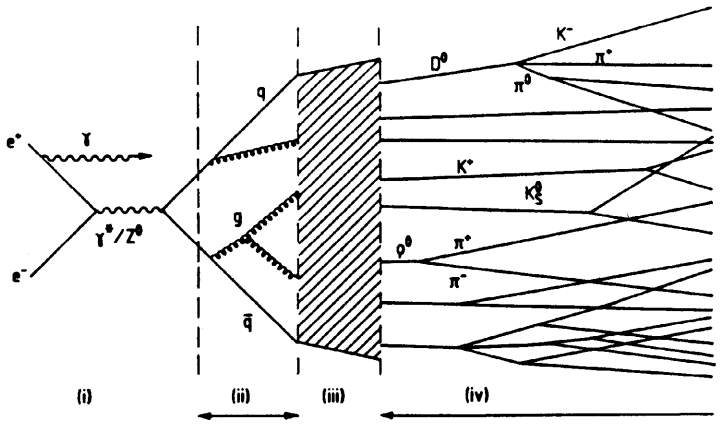

Figure 1: Schematic illustration of an e^+e^- annihilation event.

2 Physics Aspects of QCD Event Generators

The structure of a typical multihadronic event in e^+e^- annihilation is shown in Fig. 1. A few introductory paragraphs, based on this figure, follow. A number of complications are here swept under the carpet, and only covered (if at all) in the subsequent, more detailed discussion.

In a first phase, an e^+e^- pair annihilates into a virtual γ/Z^0 resonance, which produces a primary $q\bar{q}$ pair. Before the annihilation, initial state QED bremsstrahlung may occur, so that the mass of the hadronic final state is reduced from the naive value. For the precision needed for Z^0 line shape studies, also higher order (loop) corrections to the basic graph are important.

In the second phase, the initial $q\bar{q}$ pair may radiate gluons, which in their turn may radiate. While the primary production is mainly given by electroweak perturbation theory, strong perturbation theory must be used to describe this second stage. The strong coupling constant being larger than the electroweak ones, the degree of accuracy is less, in particular for soft parton emission.

In the third phase, the coloured partons fragment into a number of colourless hadrons. Although we believe this process to be given by QCD, it is not perturbatively calculable, and therefore it is an area where phenomenological models have to be invoked.

In a fourth phase, unstable hadrons decay into the experimentally observable particles. This includes everything from $\pi^0 \to \gamma\gamma$ decays to long decay chains of charm and bottom hadrons. The underlying theories are QCD and QFD, but this is again a region for which QCD predictions are primitive. The main input here comes from experimentally determined branching ratios.

It is only at this stage that detectors can be set up, to catch the produced particles, and from them reconstruct the primary hard process. This is an area not to be covered by the current report.

2.1 QED and Electroweak Aspects

The electroweak aspects of LEP experiments are covered by several theory working groups [1], and electroweak event generators appear in a separate report [2]. We here therefore only reproduce some of the main formulae actually used in QCD event generators, and comment on shortcomings.

2.1.1 Cross-section Formulae

Most QCD event generators do not include initial state radiation, and therefore events are generated at a fixed CM energy. The total cross-section is then irrelevant, and only the relative flavour composition need be known. This can be found without using particularly detailed formulae. As a rule of thumb, cross-sections obtained directly from QCD programs should therefore not be used for precision tests. Fortunately, it is usually possible to interface QCD generators with programs which give cross-sections and initial state photon radiation.

To lowest approximation, the cross-section for the process $e^+e^- \to \gamma/Z^0 \to f\bar{f}$, at given $s = E_{CM}^2$, may be written as

$$\sigma_f(s) = \frac{4\pi\alpha_{em}^2}{3s} R_f(s), \tag{1}$$

where R_f gives the ratio to the lowest order QED cross-section for the process $e^+e^- \to \mu^+\mu^-$,

$$R_f(s) = N_C R_{QCD} h_f(s). \tag{2}$$

The factor of $N_C = 3$ counts the number of colours. The R_{QCD} factor takes into account QCD loop corrections to the cross-section. For the case of five effective flavours

$$R_{QCD} \approx 1 + \frac{\alpha_S}{\pi} + 1.41\left(\frac{\alpha_S}{\pi}\right)^2 + 64.84\left(\frac{\alpha_S}{\pi}\right)^3 \tag{3}$$

in the $\overline{MS}$ renormalization scheme [5,6]. The large factor in front of the α_S^3 term has caused some worry [3]. Note that R_{QCD} does not affect the relative flavour composition, and so is of peripheral interest in event generators. Extrapolations from lower energy data to LEP gives $R_{QCD} = 1.046 \pm 0.005$ [4].

The physics specific to Z^0 production is contained in the h_f factor,

$$h_f(s) = q_e^2 q_f^2 + 2q_e v_e q_f v_f \Re\chi(s) + \left(v_e^2 + a_e^2\right)\left(v_f^2 + a_f^2\right)|\chi(s)|^2. \tag{4}$$

Here $\chi(s)$ is the ratio of Z^0 and γ propagators,

$$\chi(s) = \frac{1}{4\sin^2\theta_W(1-\sin^2\theta_W)} \frac{s}{s - m_Z^2 + i m_Z \Gamma_Z}, \tag{5}$$

and $\Re\chi(s)$ the real part of $\chi(s)$. As usual, m_Z and Γ_Z are the mass and width of the Z^0 resonance, and $\sin^2\theta_W$ the weak mixing parameter. The electric charge q_f is 2/3 for $f = u, c, t$, and $-1/3$ for $f = d, s, b$. In the standard model, the axial and vector couplings of quarks to the Z^0 are

$$\begin{aligned} a_f &= \mathrm{sign}(q_f), \\ v_f &= a_f - 4q_f \sin^2\theta_W. \end{aligned} \tag{6}$$

Corrections to the above formulae come from several sources. Quark mass effects have been neglected. Expressions also for finite quark masses are available [7], and are often used in programs. Given that the currently known quarks all are light compared to the scale of the Z^0, this is no main issue. More important are the electroweak loop corrections, which have been calculated and are available in a number of muon pair event generators [2], but generally *not* in QCD event generators.

2.1.2 Angular Orientation

While pure γ exchange gives a simple $1+\cos^2\theta$ distribution for the q (and $\overline{q}$) direction in $q\overline{q}$ events, Z^0 exchange and γ/Z^0 interference results in a forward-backward asymmetry: for unpolarized incoming e^+ and e^-

$$\frac{d\sigma}{d(\cos\theta_f)} \propto h_f(s)(1+\cos^2\theta_f) + 2h'_f(s)\cos\theta_f, \tag{7}$$

with $h_f(s)$ as given above and

$$h'_f(s) = 2q_e a_e q_f a_f \Re\chi(s) + 2v_e a_e v_f a_f |\chi(s)|^2. \tag{8}$$

Corrections to the asymmetry mainly come from higher order QCD effects, such as gluon emission. Quark mass effects also appear but, as for the cross-section, are relatively unimportant at LEP. The QCD corrections have been calculated to first order in the context of γ/Z^0 exchange [8], while second order results are available only for pure γ exchange [9,10]. The formulae are lengthy; their main effect is to smear the lowest order result, i.e. to reduce any anisotropies present in two-jet systems. These effects are usually explicitly included in matrix element programs, while only the initial $q\overline{q}$ axis is determined in parton shower programs. The subsequent shower evolution then de facto leads to a smearing of the jet axis, although not necessarily in agreement with the expectations from multijet matrix element treatments.

2.1.3 Initial and Final State QED Radiation

The most important correction to the picture above is due to initial state photon radiation. By a lowering of the effective CM energy, the cross-section may be significantly changed, downwards below the Z^0 peak and upwards above it. One of the main objectives of electroweak generators is to describe this correctly [2]. Many QCD programs do not include initial state radiation at all, and even when it is included, it is often not state of the art. It therefore makes sense to interface QCD generators to one of the electroweak generators for obtaining an accurate description. Some caution is necessary, since electroweak generators normally are set up to generate only one flavour at a time, while actually up- and down-type quarks have different line shapes, due to their different patterns of γ and Z^0 couplings.

Electroweak generators usually also include, to first non-trivial order, final state QED radiation and interference between initial and final state radiation. For lepton pair production this is unambiguous, while the physics is not well understood for quark production: contrary to leptons, quarks are not asymptotically free states, but hadronize at typical distances of a few fm. Hard final state photon emission, which takes place at short timescales, should not be affected by this. It is perfectly feasible to combine

hard photon with hard gluon emission, either in a matrix element or a parton shower picture, although most programs do not do so. It is, however, less trivial to say what phase-space region should be allowed for emission: smaller, the same, or larger than that allowed for gluons [11]. As for the interference terms, the status is even more uncertain, although plausibility arguments have been raised as to why the corrections should be small at the Z^0 peak (but not necessarily away from it) [12].

In practice, final state radiation-induced corrections to the inclusive cross-section are usually small, first, since the QED analogue of R_{QCD} is small, $R_{QED} = 1 + 3\alpha_{em}/4\pi = 1.0017$, and, second, since most final state photons are soft or collinear, and thus are buried inside jets, without affecting experimental triggers. Nevertheless, the peak of the Z^0 resonance is an interesting place to study final state radiation from quarks, since the hard photons from initial state radiation are strongly suppressed at the peak.

2.2 Perturbative QCD

The modelling of perturbative QCD is, together with the fragmentation modelling, the central objective of QCD event generators. While the hard electroweak interaction provides a description of the production of a primary $q\overline{q}$ pair, the perturbative QCD description sets out to describe the emergence of multijet events. As the CM energy is increased, hard QCD emission plays an increasingly important role, relative to fragmentation, in determining the event structure. At LEP, three-, four-, and five-jet event structures are expected to abound.

Two traditional approaches to the modelling of perturbative QCD exist. One is the matrix element method, in which Feynman diagrams are calculated, order by order. In principle, this is the correct approach, which takes into account exact kinematics, and full interference and helicity structure. The only problem is that calculations become increasingly difficult in higher orders, in particular for the loop graphs. The calculations have therefore only been carried out, in full, up to $\mathcal{O}(\alpha_S^2)$. At PETRA/PEP energies, this approach has been shown to be insufficient to account for event structures (see, however, the paragraphs on optimized Q^2 scales below), and we have strong reason to believe it will fail even more spectacularly at LEP. However, the perturbative expansion by itself will be more well-behaved at LEP, due to the smaller α_S value, so inclusive measurements may well yield more reliable results [3].

The second possible approach is the parton shower one. Here an arbitrary number of branchings of one parton into two (or more) may be put together, to yield a description of multijet events, with no explicit upper limit on the number of partons involved. This is possible since the full matrix element expressions are not used, but only approximations derived by simplifying the kinematics, and the interference and helicity structure. Parton showers are therefore expected to give a good description of the substructure of jets, but in principle the shower approach has no predictive power for the rate of well-separated jets (i.e. the 2/3/4/5-jet composition). In practice, shower programs may be patched up to describe the hard gluon emission region reasonably well. Nonetheless, these programs cannot be used for absolute α_S determinations, even if they can be used to study the running of α_S with CM energy. (Not everybody may agree with this conclusion, see e.g. the NLLjet description.)

Unfortunately, we are therefore faced with a choice between matrix elements and parton showers, a choice as delicate as that between Scylla and Charybdis – whichever

you pick, you lose some important physics aspect.

As a guess, the following is what might de facto occur. The standard generators will be of the parton showers kind, for everything from total hadronic cross-sections (i.e. detector acceptance estimates), to new particle searches, to fragmentation studies. For α_S determinations, full second order matrix elements will be compared with measures which are sensitive to the emergence of a hard third jet, but not so much to the internal structure of the three main jets. For studies of the three-gluon vertex, again matrix elements will be used, certainly four-jet ones and maybe even five-jet ones. Since the absolute jet rate will likely be wrong, either separate α_S values will be determined for each jet class, or comparisons will be on a 'per 4(5)-jet event' normalization basis.

In fact, the scenario described above may be too simplistic. It may, e.g., become necessary to combine matrix elements with parton showers in more or less contrived manners, just to have a realistic description of jet substructure when looking at four-jet events. Other approaches to studying the data may also be introduced, as discussed in the QCD report [3]. We should finally remind our readers of the experience from PETRA/PEP, where different fragmentation models turned out to yield different α_S values. The connection between perturbative and non-perturbative QCD might well hold further surprises.

2.2.1 Matrix Elements

For the discussion in this section, we will use the word 'jet' also for properties on the partonic level, almost interchangeably with the word 'parton'. However, while a 'parton' is any quark or gluon appearing in the process description, a 'jet' is one or several nearby partons, lumped together according to some jet resolution criterion (see below). The 'true' number of partons in an event is thus an ill-defined concept (and may well be infinite), while the number of jets is unique (for a given jet definition).

Three-jet matrix elements:
The Born process $e^+e^- \to q\bar{q}$ is modified in first order QCD by the probability for the q or $\bar{q}$ to radiate a gluon, i.e. by the process $e^+e^- \to q\bar{q}g$, Fig. 2a. The matrix element is conveniently given in terms of scaled energy variables in the CM frame of the event, $x_1 = 2E_q/E_{CM}$, $x_2 = 2E_{\bar{q}}/E_{CM}$, and $x_3 = 2E_g/E_{CM}$, i.e. $x_1 + x_2 + x_3 = 2$. For massless quarks the matrix element reads [13]

$$\frac{d\sigma}{dx_1 dx_2} = \sigma_0 \frac{\alpha_S}{2\pi} C_F \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}, \tag{9}$$

where σ_0 is the lowest order cross-section, $C_F = 4/3$ is the appropriate colour factor, and the kinematically allowed region is $0 \le x_i \le 1, i = 1,2,3$. By kinematics, the x_k variable for parton k is related to the invariant mass m_{ij} of the remaining two partons i and j by $y_{ij} = m_{ij}^2/E_{CM}^2 = 1 - x_k$.

Jet resolution criteria:
In order to separate two-jets from three-jets, it is useful to introduce jet resolution parameters, either (ϵ, δ) or y.

- (ϵ, δ) : a three-parton configuration is called a two-jet event if $\min(x_i) < \epsilon$ or if $\min(\theta_{ij}) < \delta$ (where θij is the angle between partons i and j in the CM frame of the event).
- y : a three-parton configuration is called a two-jet event if $\min(y_{ij}) = \min(m_{ij}^2/E_{CM}^2) < y$.

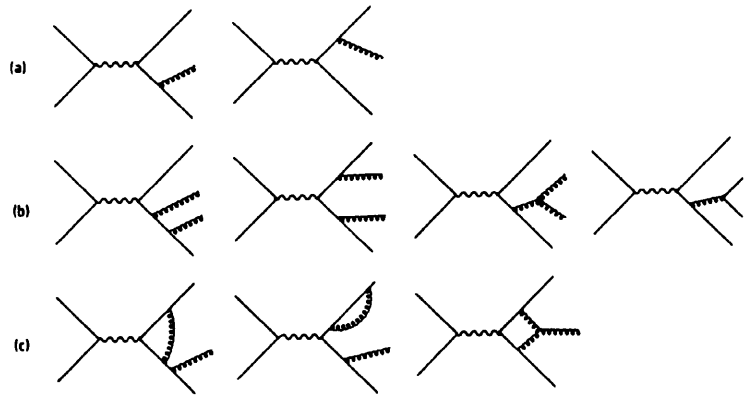

Figure 2: Feynman graphs for three- and four-jet production.
(a) The two graphs which contribute to three-jet production.
(b) A few of the graphs which contribute to four-jet production (the ones left out can be obtained by symmetry).
(c) A few of the loop (vertex and propagator) graphs which contribute to three-jet production in second order.

In the discussion which follows, we will mainly refer to the y cut case, since this is the simpler one.

The cross-section in eq. (9) diverges for $x_1 \to 1$ or $x_2 \to 1$ but, when first order propagator and vertex corrections are included, a corresponding singularity with opposite sign appears in the $q\bar{q}$ cross-section, so that the total cross-section is finite. In analytical calculations, the average value of any well-behaved quantity $\mathcal{Q}$ can therefore be calculated as

$$\langle \mathcal{Q} \rangle = \frac{1}{\sigma_{tot}} \lim_{y \to 0} \left(\mathcal{Q}(2parton)\sigma_{2parton}(y) + \int_{y_{ij}>y} \mathcal{Q}(x_1, x_2) \frac{d\sigma_{3parton}}{dx_1 dx_2} dx_1 dx_2 \right), \tag{10}$$

where any explicit y dependence disappears in the limit $y \to 0$.

In Monte Carlo programs, it is not possible to work with a negative total two-jet rate, and so it is necessary to introduce a fixed non-vanishing y cut-off in the three-jet. phase-space. Experimentally, there is evidence for the need of a low y cut-off, i.e. a large three-jet rate. For LEP applications, the recommended value is $y = 0.01$, which is about as far down as one can go and still retain a positive two-jet rate. With $\alpha_S = 0.12$, in full second order QCD, the $2:3:4$ jet composition is then approximately $11\% : 77\% : 12\%$. Note, however, that initial state QED radiation may occasionally lower the CM energy significantly, i.e. increase α_S, and thereby bring the three-jet fraction above unity if y is kept fixed at 0.01 also in those events. (At PETRA/PEP, y values slightly above 0.01 were needed.)

The strong coupling constant:
The strong coupling constant α_S is in first order given by

$$\alpha_S(Q^2) = \frac{12\pi}{(33-2n_f)\ln(Q^2/\Lambda^2)}. \tag{11}$$

The number of flavours n_f is 5 for LEP applications, and so the Λ value determined is Λ_5, while e.g. most deep inelastic scattering studies refer to Λ_4, the energies for these experiments being below the bottom threshold.

In higher orders the result depends on the renormalization scheme; we will use $\overline{MS}$ throughout. In addition, several approximations have been in use: one can define the higher order contributions in terms of the first order coupling constant, i.e. by an expansion in $1/\ln(Q^2/\Lambda^2)$ [14],

$$\alpha_S(Q^2) = \frac{12\pi}{(33-2n_f)\ln(Q^2/\Lambda^2_{\overline{MS}})}\left[1 - 6\frac{153-19n_f}{(33-2n_f)^2}\frac{\ln(\ln(Q^2/\Lambda^2_{\overline{MS}}))}{\ln(Q^2/\Lambda^2_{\overline{MS}})}\right], \tag{12}$$

or one can sum the leading logarithms in this expansion, i.e. terms proportional to $\alpha_S^n \ln^n(Q^2/\Lambda^2)$, to obtain

$$\alpha_S(Q^2) = \frac{12\pi}{(33-2n_f)\ln(Q^2/\Lambda^2_{\overline{MS}}) + 6\frac{153-19n_f}{33-2n_f}\ln(\ln(Q^2/\Lambda^2_{\overline{MS}}))}. \tag{13}$$

Both definitions have been widely used. Numerically they give somewhat different results, e.g., for $\alpha_S = 0.15$ and $n_f = 5$, the Λ values from the two definitions differ by 14%. If one includes the next higher order term in eq. (12), the difference with eq. (13) becomes almost negligible.

Instead of using a series expansion in α_S, one can solve the renormalization group equation, which yields [6,15]

$$\ln\frac{Q^2}{\Lambda^2_{\overline{MS}}} = \frac{12\pi}{(33-2n_f)\alpha_S} - 6\frac{153-19n_f}{(33-2n_f)^2}\ln\left[\frac{12\pi}{(33-2n_f)\alpha_S} + 6\frac{153-19n_f}{(33-2n_f)^2}\right] \tag{14}$$

The numerical solution to this formula yields Λ values in between the ones from eqs. (12) and (13). The third order term, not included above, is numerically insignificant, which is a nice feature compared with the approximate series expansion in α_S, where the third order term changes Λ by about 7%.

Four-jet matrix elements:
Two new event types are added in second order QCD, $e^+e^- \to q\bar{q}gg$ and $e^+e^- \to q\bar{q}q'\bar{q}'$, Fig. 2b. The four-jet cross-section has been calculated by several groups [9,16,17,18], which agree on the result. The formulae are too lengthy to be quoted here. In one of the calculations [9], quark masses were explicitly included. The original calculations were for the pure γ exchange case; recently it has been pointed out [19] that an additional contribution to the $e^+e^- \to q\bar{q}q'\bar{q}'$ cross-section arises from the axial part of the Z^0. This term is not included in any program so far, but fortunately it is finite and small.

Second order three-jet corrections:
As for first order, a full second order calculation consists both of real parton emission terms, and of vertex and propagator corrections, Fig. 2c. These modify the three-jet and two-jet cross-sections. Although there was some initial confusion, everybody

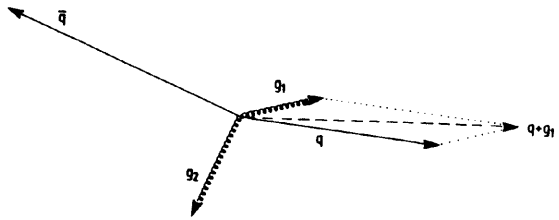

Figure 3: A four-parton event, with two nearby partons that are to be recombined. Construction of momentum sum $q+g_2$ is shown, but that gives a composite parton with non-zero mass.

soon agreed on the size of the loop corrections [17,20,21]. In analytic calculations, the procedure of eq. (10), suitably expanded, can therefore be used unambiguously for a well-behaved variable. The second order corrections will also modify the first order α_S expression, as described above; this is mainly of interest in relating an experimentally measured value of α_S to a Λ value.

For Monte Carlo event simulation, it is again necessary to impose some finite jet resolution criterion. This means that four-parton events which fail the cuts should be reassigned either to the three-jet or to the two-jet event class, see Fig. 3. It is this area which has caused quite a lot of confusion in the past [22], and where full agreement does not exist. Most likely, agreement will never be reached, since there are indeed ambiguous points in the procedure, related to uncertainties on the theoretical side. This is illustrated in the following paragraphs.

For the y-cut case, any two partons with an invariant mass $m^2_{ij} < yE^2_{cm}$ should be recombined into one. If the four-momenta are simply added, the sum will correspond to a parton with a positive mass, namely the original m_{ij}. The loop corrections are given in terms of final massless partons, however. In order to perform the (partial) cancellation between the four-parton real and the three-parton virtual contributions, it is therefore necessary to get rid of the bothersome mass in the four-parton states. Several procedures are used in practice; the following two are probably the most frequently used.

- The $\overline{p}$ recombination scheme: keep the constructed three-momentum sum $\overline{p}_{ij} = \overline{p}_i + \overline{p}_j$ (in the CM frame of the event), and redefine the energy of the recombined parton as being $E_{ij} = |\overline{p}_i + \overline{p}_j|$. Since $E_{ij} < E_i + E_j$, the total CM energy of the event has been reduced, which can be compensated by rescaling the three four-momenta by a common factor.
- The E recombination scheme: require $E_{ij} = E_i + E_j$. The three variables x_1, x_2 and x_3 are then easily obtained, but the three-momenta have to be modified in a non-trivial manner to keep momentum conserved.

The E scheme obviously gives more energy to the recombined jet than does the $\overline{p}$ one. Since the typical situation is that the recombined jet still is the lowest-energy one (two soft gluons recombined into a medium soft one), the E scheme gives more three-jetlike topologies than the $\overline{p}$ one, and so has a larger second order correction to the three-jet rate. Differences need not be negligible.

Within each scheme, a number of lesser points remain to be dealt with, in particular what to do if a recombination of a nearby parton pair were to give an event with a non-$q\bar{q}g$ flavour structure.

If the situation is complicated enough for the y cut, it is even worse for the (ϵ, δ) alternative. Separate recipes need to be specified for what to do when a parton fails the ϵ cut, when a pair fails the δ cut, and when both happen in the same event. In particular, for the ϵ cut, one could either discard the soft parton outright, or choose between several possible recombination algorithms.

The physical spread, in terms of final effective three-jet rate, i.e. in terms of α_S values found for given data, has therefore always been larger in the (ϵ, δ) alternative. (If soft partons always are recombined, like in the y cut alternative, the two approaches become more comparable, however.) This is the reason why, in section 3, we will concentrate on programs based on y cuts. This is not a completely unambiguous decision: the y cut-off approach is less well suited for giving the second order corrections to the energy-energy correlation and its asymmetry, for example.

Optimized perturbation theory:
Theoretically, it turns out that the second order corrections to the three-jet rate are large. It is therefore not unreasonable to expect large third order corrections to the four-jet rate. Indeed, the experimental four-jet rate is much larger than second order predicts (when fragmentation effects have been folded in), if α_S is determined based on the three-jet rate [23].

The only consistent way to resolve this issue is to go ahead and calculate the full next order. This is a tough task, however, so people have looked at possible shortcuts. For example, one can try to minimize the higher order contributions by a suitable choice of the renormalization scale [15] — 'optimized perturbation theory'. This is equivalent to a different choice for the Q^2 scale in α_S, a scale which is not unambiguous anyway. Indeed the standard value $Q^2 = s = E_{CM}^2$ is larger than the natural physical scale of gluon emission in events, given that most gluons are fairly soft. One could therefore pick another scale, $Q^2 = y's$, with $y' < 1$. The $\mathcal{O}(\alpha_S)$ three-jet rate would be increased by such a scale change, and so would the number of four-jet events, including those which collapse into three-jet ones. The loop corrections depend on the Q^2 scale, however, and compensate the changes above by giving a larger negative contribution to the three-jet rate.

Implementing a different scale in a Monte Carlo program can be done easily as follows. Suppose the three-jet rate is given in a certain renormalization scheme and for a given Q^2 scale [15,24]

$$R_3 = r_1\alpha_S + r_2\alpha_S^2 + \mathcal{O}(\alpha_S^3). \tag{15}$$

When the coupling is chosen at a different scale, $Q'^2 = y'Q^2$,

$$R_3' = r_1'\alpha_S' + r_2'\alpha_S'^2 + \mathcal{O}(\alpha_S'^3). \tag{16}$$

If one neglects the terms of $\mathcal{O}(\alpha_S^3)$, the two should be the same,

$$R_3' - R_3 = dR_3 = r_1 d\alpha_S + \alpha_S dr_1 + \alpha_S^2 dr_2 = 0 \tag{17}$$

(note that $d\alpha_S$ is of $\mathcal{O}(\alpha_S^2)$). Since each of the powers of α_S have to vanish separately, $r_1' = r_1$, and (after calculating $d\alpha_S$ from eq. (11))

$$r_2' = r_2 + r_1\frac{33 - 2n_f}{12\pi}\ln y'. \tag{18}$$

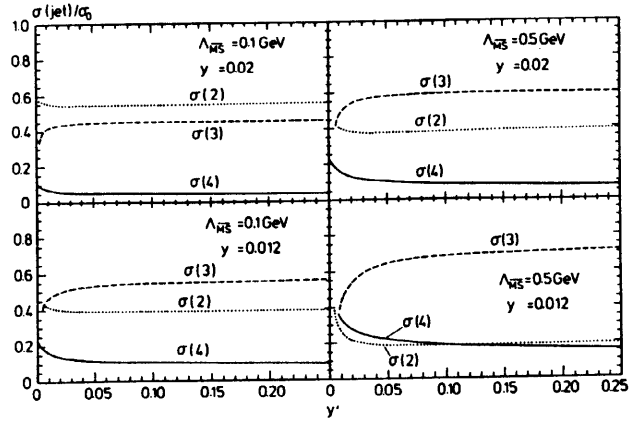

Figure 4: Jet cross-sections as a function of the scale parameter y', for different values of the QCD scale (0.1 and 0.5 GeV) and the cut-off parameter y (0.01 and 0.03).

The procedure above, while given for the total three-jet rate, could equally well be carried out for the differential cross-section in a phase-space point (x_1, x_2), with the same result.

Since we only have the Born term for four-jets, here the effects of a scale change come only from the change in the coupling constant. Finally, the two-jet cross-section can still be calculated from the difference between the total cross-section and the three- and four-jet cross-sections. Fig. 4 shows this procedure applied to the cross-sections in JETSET.

In a few studies [25,26,27,28], it has been concluded that a y' value around or even somewhat below the y matrix element cut-off value gives significantly improved agreement with the data. In particular, the four-jet rate is directly proportional to α_S^2, and so is larger for smaller y'. A manipulation of the y' scale is thus one way to change the relative rate of four- to three-jet production. As a by-product, it is also possible to reduce the Λ scale by about a factor of 2, which brings it in better agreement with deep inelastic scattering values. (A similar conclusion holds for the energy-energy correlation asymmetry, for which the higher order corrections are rather small [27]).

The success of describing the jet rates should not hide the fact that one is dabbling in (educated, hopefully) guesswork, and that any conclusions based on this method have to be taken with a pinch of salt. (For example, the exact third order expression for R_{QCD} is not in agreement with expectations of optimized perturbation theory applied to the second order expression.) However, in principle one should be better off at LEP than at PETRA, see [3].

Higher order matrix elements:
In the last year, several groups have calculated the five-jet Born term [29,30]. This is certainly impressive, and could be useful in studying e.g. the three-gluon vertex. It is very likely that the results will not be included in standard Monte Carlo programs, for a number of reasons. First, the formulae are lengthy; generation of unweighted events is therefore going to be slow. Second, the actual rate of five-jet production is small, even

when increased by the use of an optimized scale as described above. Third, inclusion of five-jets will not help at all for α_S determinations so long as loop corrections to the same order are not available. Furthermore, given the large third order corrections to R_{QCD}, eq. 3, there is no reason to expect those corrections to be small.

The techniques developed not only make it possible to calculate five-jet matrix elements, but also processes of the type $q\bar{q}(ng), n \geq 0$, i.e. with an arbitrary number of gluons [30]. Matrix element evaluation is slow for many gluons, but simpler approximate expressions exist [31], which preserve the full pole structure, and numerically seem to agree with the exact results on the 10% level. These could come in very handy for topological studies.

Energy dependence:
Finally, one should note that fragmentation parameters determined at lower energies, for models based on matrix elements, are not expected to give a good description at LEP energies, since typically a cut-off more or less fixed in y is used. Therefore the actual minimum invariant mass between partons increases with CM energy. This means that an increasingly large region of soft to medium soft gluons is excluded, not because we believe this is the correct thing to do, but because including them would give a negative two-jet (or three-jet) probability, which is technically excluded. To compensate somewhat for the effects of neglected gluon emission, it is therefore expected that one must use energy-dependent fragmentation parameters, by which jets become softer and broader at higher energies.

2.2.2 Parton Showers

The parton shower picture is derived within the framework of the leading logarithm approximation, the LLA. In this picture, only the leading terms in the perturbative expansion are kept and, where need be, resummed. Subleading corrections, which are down in order by factors of $\ln Q^2$ or $\ln z$ ($\ln(1-z)$), or by powers of $1/Q^2$, are thus neglected. Different schemes have been devised for taking into account some subleading corrections; they appear under names like MLLA (M for modified), DLLA (D for double) or NLLA (N for next-to-) [32]. The overall theoretical picture is rather encouraging: there is reason to believe that neglected subleading effects are small, and the predictive power of this approach is increasing year by year.

Phenomenologically, the main reason for the LLA success is its ability to be formulated in terms of a probabilistic picture, suitable for event generation. The approach is based on simplifications in kinematical variables, however, so the predictive power for hard, wide-angle parton emission remains limited.

The evolution equations:
Most parton shower algorithms are based on an iterative use of the basic branchings $q \to qg$, $g \to gg$, and $g \to q\bar{q}$, Fig. 5. A probabilistic picture is used to describe these branchings, as follows.

The probability $\mathcal{P}$ that a branching $a \to bc$ will take place during a small change $dt = dQ^2_{evol}/Q^2_{evol}$ of the evolution parameter $t = \ln(Q^2_{evol}/\Lambda^2)$ is given by the Altarelli-

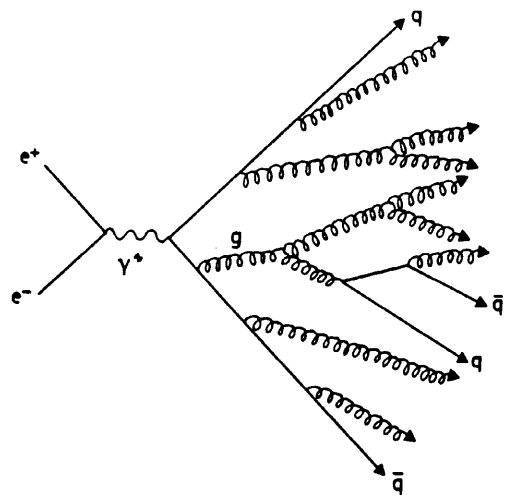

Figure 5: Schematic picture of parton shower evolution in e^+e^- events.

Parisi equations [33]

$$\frac{d\mathcal{P}_{a\to bc}}{dt} = \int dz \frac{\alpha_S(Q^2)}{2\pi} P_{a\to bc}(z). \tag{19}$$

Summation over all possible final state flavour combinations b and c is implied for gluon branchings. The $P_{a\to bc}(z)$ are the Altarelli-Parisi splitting kernels

$$\begin{aligned} P_{q\to qg}(z) &= C_F \frac{1+z^2}{1-z}, \\ P_{g\to gg}(z) &= N_C \frac{(1-z(1-z))^2}{z(1-z)}, \\ P_{g\to q\bar{q}}(z) &= T_R(z^2+(1-z)^2), \end{aligned} \tag{20}$$

with $C_F = 4/3$, $N_C = 3$, and $T_R = n_f/2$, i.e. T_R receives a contribution of 1/2 for each allowed $q\bar{q}$ flavour. The z variable specifies the sharing of four-momentum between the daughters, with daughter b taking fraction z and c taking $1-z$. Usually the first order α_S is used, eq. (11); as we shall see, the Q^2 scale of α_S need not agree with the evolution scale Q^2_{evol}.

Persons familiar with analytical calculations may wonder why the '+ prescriptions' and $\delta(1-z)$ terms of the splitting kernels in eq. (20) are missing. These complications fulfil the task of ensuring flavour and energy conservation in the analytical equations. The corresponding problem is solved trivially in Monte Carlo programs, where the shower evolution is explicitly traced, and flavour and four-momentum are conserved at each branching. The legacy left is the need to introduce a cut-off on the range of the z integral in the Altarelli-Parisi equations, so as to avoid the singular regions corresponding to excessive production of very soft gluons. Typically, this is achieved by introducing an effective (fictitious) gluon mass (in programs usually denoted by Q_0 or $Q_0/2$).

Also note that $P_{g\to gg}(z)$ is given here with a factor 3 in front, while it is sometimes shown with 6. The difference of a factor of two comes from either considering the number of gluons that branch or the number of gluons that are produced.

The Sudakov form factor:
Starting at the maximum allowed virtuality t_{max} for parton a, the t parameter may be successively degraded. (This does not mean that an individual parton runs through a range of t values: each parton in the end is associated with a fixed t value, and the evolution procedure is just a way of picking that value. It is only the ensemble of partons in many events that evolve continuously with t, cf. the concept of structure functions.) The probability that no branching occurs during a small range of t values, δt, is given by $(1-\delta t\, d\mathcal{P}/dt)$. When summed over many small intervals, the no-emission probability exponentiates

$$\mathcal{P}_{no-emission}(t_{max},t) = \exp\left(-\int_t^{t_{max}} dt' \frac{d\mathcal{P}_{a\to bc}}{dt'}\right). \tag{21}$$

Thus the probability for a branching at a given t is the naive probability for a branching, eq. (19), multiplied by the probability that a branching has not already taken place, eq. (21). This is nothing but the exponential decay law of radioactive decays, with a t-dependent decay probability.

It is customary to introduce the Sudakov form factor, i.e. the probability that a parton starting from a maximum virtuality t will reach the fixed lower cut-off t_{min} (related to the effective gluon mass Q_0) without branching

$$S_a(t) = \exp\left(-\int_{t_{min}}^{t} dt' \int_{z_{min}(t')}^{z_{max}(t')} dz \frac{\alpha_S(Q^2)}{2\pi} P_{a\to bc}(z)\right). \tag{22}$$

The no-emission probability above is then just $S_a(t_{max})/S_a(t)$. Since the Sudakov form factor only depends on one parameter, t, it is easy to pretabulate, for each flavour a, at the beginning of a Monte Carlo run. Many programs use this as part of the generation strategy.

With R a random number uniformly distributed between 0 and 1, the t value for a branching may be found by solving the equation

$$\mathcal{P}_{no-emission}(t_{max},t) = R, \tag{23}$$

i.e.

$$S_a(t) = \frac{S_a(t_{max})}{R}. \tag{24}$$

The products b and c may be allowed to branch in their turn, and so on, giving a treelike structure. The branching of a given parton is stopped whenever the evolution parameter is below t_{min}, i.e. when an $R < S_a(t_{max})$ is chosen.

Coherence:
Very valuable input for model builders is provided by the theoretical studies of corrections beyond leading log, like coherence effects [34,35]. These come in two kinds.

- The intrajet coherence phenomenon is responsible for a decrease of the amount of soft gluon emission inside jets. It has been shown that an ordering in terms of a decreasing emission angle takes into account the bulk of soft gluon interference effects. Programs which contain angular ordering are loosely called coherent showers, while those without are called conventional ones.

- The interjet coherence phenomenon, responsible for the flow of particles in between jets, with constructive or destructive interference depending on colour configuration ('colour drag phenomena'), cf. [35]. This form of coherence is not a direct consequence of the ordering of (polar) emission angles mentioned above, but rather requires that azimuthal angles of branchings be properly distributed.

Further, theoretical studies of loop corrections [32] strongly suggest the use of a scale $Q^2 = z(1-z)m_a^2 \approx p_T^2$ as the argument of α_S, i.e. the scale is set by the transverse momentum of a branching, rather than by the mass of the decaying parton, as might naively have been expected.

On theoretical grounds, there is no doubt that coherence phenomena have to be present in nature. However, the enthusiasm over the coherence phenomenon has led to the emergence of a folklore, which is sometimes too simplistic. One myth is that coherence is already experimentally proved by the 'string effect' of JADE [36]. At the energies so far explored, all the effects could be fully explained (and were actually predicted [37]) based on non-perturbative dynamics. Only with increased energy, and hence an increased number of soft gluons, would perturbative effects be expected to dominate. Therefore coherence effects could be a main subject for LEP QCD studies; see also section 5.7.

Another piece of folklore is that coherent parton showers predict a slower growth of (parton) multiplicity with energy than do conventional ones. Specifically, in coherent showers, the multiplicity is expected to have an energy dependence (neglecting subleading effects) of the form [38]

$$\langle n(t)\rangle \propto t^c \exp\left(\sqrt{\frac{72}{33-2n_f}t}\right), \quad c = -\frac{1}{4}\left(1+\frac{5}{9}\frac{8n_f}{33-2n_f}\right) \approx -\frac{1}{2}, \tag{25}$$

whereas an additional factor of 2 should appear in the exponential in the conventional case (also c would be modified). The expected difference comes from comparing the allowed range of z values for branchings in the two approaches. This range should naively be smaller in coherent algorithms. However, in several conventional shower programs (like the one in [39]), additional kinematics-related cuts on z are introduced. These cuts are not equivalent to those of coherent showers, but are still severe enough to give a slow growth at current energies. The expected difference in multiplicity growth will then be visible only when α_S is truly small.

Shower programs, the basic approaches:
A wide selection of shower algorithms have been developed (those in section 4, and more [39,40,41]), which mainly differ in the interpretation of the variables t (or Q^2_{evol}), Q^2 and z. Many of the variations are formally of a subleading character, and therefore are not constrained by theoretical leading log analyses, while others indeed imply quite different physics, at least at higher energies. The actual algorithms used are sufficiently different that it would be difficult here to give an overview of them all. For details we instead refer to the program descriptions in section 4, and to the QCD theory report [3].

While z generically expresses the sharing of energy and momentum between the two daughters, the exact definition varies from program to program. It is possible to define z in terms of energy E, in terms of light-cone momentum $E+p_L$, in terms of $E+|p|$, or in terms of some other combination of energy and momentum. The definition chosen

affects the allowed region of z values, i.e. the $z_{min}(t)$ and $z_{max}(t)$ functions of eq. (22). The introduction of angular ordering actually constrains the allowed z range, compared to the kinematically allowed one. Since the details of the z interpretation are most significant at low z values, angular ordering as a 'fringe benefit' introduces a reduced dependence on the z choice made.

One should note that none of the z definitions given above are fully Lorentz covariant, and hence neither are most shower algorithms. From a theoretical point of view, this is related to the need to fix a gauge before the branching probabilities can be evaluated (it is e.g. possible to have an algorithm where only one of the initial two partons radiate). The general belief is that this is a rather minor point, with no experimentally observable consequences, but it is worth bearing in mind.

The best known coherent shower algorithm is probably the Marchesini-Webber one, where angular ordering is built in from the onset. Here $Q^2_{evol} = E^2\xi$ with $\xi \approx 1 - \cos\theta$. E is the energy of the branching parton and θ the opening angle between its decay products. At each branching, an actual Q^2_{evol} is selected from tabulated Sudakov form factors, using eq. (24). The Q^2_{evol} defines the allowed z range, within which a value is picked. Q^2_{evol} gives the ξ_a value of the branching, z the sharing of energy, $E_b = zE_a, E_c = (1-z)E_a$. Because of the requirement of ordering in angle, the maximum Q^2_{evol} scale of the two daughters is now given by $E_b^2\xi_a$ and $E_c^2\xi_a$, respectively. These daughters may be degraded in their turn, to find $\xi_b < \xi_a$ and $\xi_c < \xi_a$, etc.

The Marchesini-Webber algorithm is different from others in several respects. One is that parton masses are not defined during the evolution stage. Only afterwards are parton masses constructed from the final on-shell partons backwards to the shower initiator, by using the opening angle variable, and only after that is the actual kinematics of the shower found. Most other programs construct the kinematics in parallel with the shower evolution proper. Another difference is that the Marchesini-Webber algorithm makes use of an evolution variable which explicitly involves angles, and which thereby automatically incorporates angular ordering.

In a more typical algorithm, such as those implemented in JETSET and CALTECH-II, the evolution variable is defined to be $Q^2_{evol} = m_a^2$, i.e. the mass of the decaying parton. This means that angular ordering is not automatically included, but must be imposed as an additional constraint on the combination of m^2 and z values which are allowed for a particular branching, given the m^2 and z values of the preceding branching. Even if the actual phase-space in the end comes out the same as in HERWIG (which is true to the level of approximation to which we can trust the leading log picture anyway), and even if the naive probability for a given branching is the same, the end result will be different due to the difference in the Sudakov form factor, which reflects the choice of evolution variable. In the HERWIG case, emission angles close to the preceding ones will be slightly favoured, since they will be considered 'first' and therefore correspond to a negligible Sudakov damping, while m^2 evolution programs for the same reason will favour large masses (all other things being equal).

Shower programs, possible improvements:

Since showers are so important for phenomenological studies at higher energies, a steady evolution is taking place in the field. Three main approaches can be distinguished for this improvement. One is to refine the standard leading log picture as such, by including further corrections and effects. This is the line of HERWIG and other programs.

Another is to go beyond leading log, and to also include higher order effects, as is done in NLLjet. The third is to strive for an alternative (but equivalent) formulation of the shower process, where a number of non-trivial effects are included automatically. This is exemplified by ARIADNE. A few words about each of these.

One way to improve on the basic leading log picture is to include non-isotropic azimuthal angles. This is particularly well explored by HERWIG, in which the following three sources of anisotropy are included:

- Gluon polarization, leading to a correlation between production and decay planes of a gluon.
- Gluon polarization, also giving correlations between non-adjacent branchings, i.e. of the 'Bell inequality' type [42].
- Soft gluon interference, related to the fact that the soft gluon emission probability is actually obtained by summing emission amplitudes from a number of separate hard partons.

Another possible improvement, found e.g. in JETSET, is to constrain the first branchings of the shower to agree with the explicit three-jet matrix element form. This is an attempt to modify the shower formalism in the region where the kinematical approximations involved are known to be least reliable, while still preserving concepts of the LLA such as the Sudakov form factor and the $Q^2 \approx p_T^2$ argument in α_S. To date, it has not been shown whether this feature helps improve agreement with data or not. It certainly does not improve the ability to use the program for α_S determinations, since so many uncertainties remain anyway.

As the name suggests, NLLjet is a program which tries to go beyond leading order, to Next-to-Leading-Logs. The most apparent consequence is the introduction of $1 \to 3$ parton branchings: $q \to qgg, q \to qq'\bar{q}', g \to ggg$, and $g \to gq\bar{q}$. The ordinary $1 \to 2$ branchings are also modified, in analogy with the second order three-jet matrix element modifications. Since loop graphs are explicitly involved in calculating the corrections, the Q^2 scale definition is under better control. From that point of view, NLLjet should be superior for Λ determinations. Unfortunately, as for leading logs, the whole approach is really only valid for collinear kinematics. A number of additional assumptions are needed for extrapolations to the interesting region of well separated jets. One past study demonstrated rather large differences between the next-to-leading-log four-jet topologies and the explicit four-jet matrix elements [43]. While these results do not necessarily carry over to NLLjet, the question of non-collinear kinematics remains a crucial one.

An alternative to parton shower algorithms is the dipole formulation of ARIADNE, suggested by the Leningrad group [44], and studied in detail by the Lund people. The picture is based on identifying the string pieces between partons with colour dipoles or colour antennae, so that the emission of a gluon corresponds to the breaking of a dipole into two. This breaking is simple and well-defined in the rest frame of a dipole, and yet it automatically includes angular ordering and non-trivial azimuthal effects when the boost back to the overall CM frame is taken into account. The picture has a number of appealing features, e.g. an explicitly Lorentz covariant formulation, and may well come closer to describing nature as it really is than do the other algorithms. This does not mean it is a unique recipe. While angular ordering is built in, the transverse momentum of a dipole branching is a priori not required to be smaller than the p_T of the branching in which it was produced; the model is rather sensitive to what is assumed on this

count. Additional degrees of freedom in branchings come from the angular orientation of daughter dipoles with respect to the mother one.

2.3 Fragmentation

The fragmentation process has yet to be understood from first principles, starting from the QCD Lagrangian. This has left the way clear for the development of a number of different phenomenological models. Being models, none of them can lay claims to being 'correct'. The best one can aim for is a good representation of existing data, plus a predictive power for properties not yet studied or results at higher energies.

All existing models are of a probabilistic and iterative nature. This means that the fragmentation process as a whole is described in terms of one (or a few) simple underlying branchings, of the type jet → hadron + remainder-jet, string → hadron + remainder-string, cluster → hadron + hadron, or cluster → cluster + cluster. At each branching, probabilistic rules are given for the production of new flavours, and for the sharing of energy and momentum between the products.

Three main schools are usually distinguished, independent fragmentation (IF), string fragmentation (SF) and cluster fragmentation (CF) [45]. These need not be mutually exclusive; it is possible to have models which contain both cluster and string aspects, or models which interpolate between independent and string fragmentation.

While the evolution of fragmentation models was rapid in the early eighties, no really new algorithms have been introduced in the last five years, and only a modest amount of refinement of the existing approaches has been performed. New concepts, like local parton-hadron duality, and new experimental features, like intermittency, have recently led to a resurgence of fragmentation studies outside the framework of the existing programs. While very interesting, these studies have not yet led to fully-fledged programs of the kind needed to describe LEP events in full, and maybe they never will. A few examples of new ideas will appear later in this report, as a reminder for users to keep an open mind, given that nobody knows the ultimate truth.

2.3.1 Independent Fragmentation

The independent fragmentation (IF) approach dates back to the early seventies [46], and gained widespread popularity with the Field-Feynman paper [47]. Subsequently, IF was the basis for two programs widely used in the early PETRA/PEP days, the Hoyer *et. al.* [48] and the Ali *et. al.* [49] programs. These programs have not been updated since their conception, and are therefore not well suited for LEP physics studies. Among the programs covered in this report, EURODEC is a descendant of the Ali *et. al.* program, COJETS uses a special IF approach, and JETSET has as (non-default) options a wide selection of independent fragmentation algorithms.

In the IF approach, it is assumed that the fragmentation of any system of partons can be described as an incoherent sum of independent fragmentation procedures for each parton separately. The process is to be carried out in the overall CM frame of the jet system, with each jet fragmentation axis given by the direction of motion of the corresponding parton in that frame.

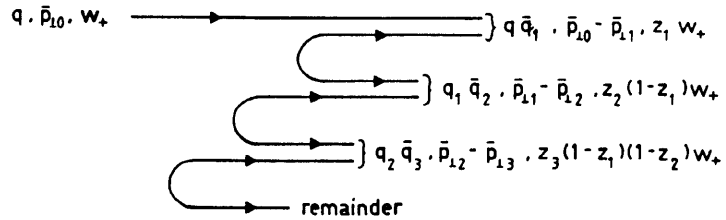

Figure 6: The iterative ansatz for flavour, transverse momentum, and light-cone energy-momentum ($W_+ = E + p_L$) fraction.

The iterative ansatz:
Both in independent and string fragmentation, the fragmentation of a jet can be described iteratively, as follows. Assume that a quark is kicked out by some hard interaction, carrying a well-defined amount of energy and momentum. This quark jet q is split into a hadron $q\bar{q}_1$ and a remainder-jet q_1, essentially collinear with each other. The sharing of energy and momentum is given by some probability distribution $f(z)$, where z is the fraction taken by the hadron, leaving $1 - z$ for the remainder-jet. The remainder-jet is assumed to be just a scaled-down version of the original jet, in an average sense. The process of splitting off a hadron can therefore be iterated, to yield a sequence of hadrons, Fig. 6. In particular, the function $f(z)$ is assumed to be the same at each step, i.e. independent of remaining energy. If z is interpreted as the fraction of the jet $E + p_L$, i.e. energy plus longitudinal momentum w.r.t. the jet axis, this leads to a flat central rapidity plateau dn/dy for a large initial energy.

The normal z interpretation means that a choice of a z value close to 0 corresponds to a particle moving backwards, i.e. with $p_L < 0$. It makes sense to allow only the production of particles with $p_L > 0$, but to explicitly constrain z accordingly would destroy longitudinal invariance. The most straightforward way out is to allow all z values but discard hadrons with $p_L < 0$. Flavour, transverse momentum and $E + p_L$ carried by these hadrons are 'lost' for the forward jet. The average energy of the final jet comes out roughly right this way, with a spread of 1 - 2 GeV around the mean. The jet longitudinal momentum is decreased, however, since the jet acquires an effective mass during the fragmentation procedure. For a two-jet event this is as it should be, at least on average, because also the momentum of the compensating opposite-side parton is decreased.

A number of different fragmentation functions $f(z)$ have been proposed.

- The Field-Feynman parametrization,

$$f(z) = 1 - a + 3a(1 - z)^2, \tag{26}$$

  with default value $a = 0.77$, is frequently used for ordinary hadrons.
- Since there are indications that the shape above is too strongly peaked at $z = 0$, instead a shape like

$$f(z) = (1 + c)(1 - z)^c \tag{27}$$

  may be used.

- Charm and bottom data clearly indicate the need for a harder fragmentation function for heavy flavours, see section 5.6. The best known of these is the Peterson *et. al.* formula [50]

$$f(z) \propto \left[z \left(1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z} \right)^2 \right]^{-1}, \tag{28}$$

where ϵ_Q is a free parameter, expected to scale between flavours like $\epsilon_Q \propto 1/m_Q^2$.

Each splitting corresponds to the production of a $q'\overline{q}'$ pair. The flavour of this pair is chosen at random according to the relative probabilities $u\overline{u} : d\overline{d} : s\overline{s} = 1 : 1 : \gamma_s$, typically with the parameter $\gamma_s = 0.3$. It is also possible to allow for the production of diquark-antidiquark pairs, with corresponding probabilities to be given. While the flavour content of each hadron is thus uniquely specified by the flavours of two adjacent splittings, it is still possible to generate different spin and orbital angular momentum. Normally only the lowest-lying pseudocalar and vector meson multiplets are included, with relative probability given by a parameter. For diagonal flavour combinations, like $u\overline{u}$, it is necessary to specify the mixing between $\pi^0 - \eta - \eta'$ and $\rho^0 - \omega - \phi$.

In addition to local flavour conservation in $q'\overline{q}'$ splittings, it is also assumed that transverse momentum is locally conserved, i.e. the net p_T of the $q'\overline{q}'$ pair as a whole is assumed to be vanishing. The p_T of the q is taken to be a Gaussian in the two transverse degrees of freedom separately, with the transverse momentum of a hadron obtained by the sum of constituent quark transverse momenta.

Gluon jets:
Within the IF framework, there is no unique recipe for how gluon jet fragmentation should be handled. One possibility is to treat it exactly like a quark jet, with the initial quark flavour chosen at random among $u, \overline{u}, d, \overline{d}, s$ and $\overline{s}$, including the ordinary s quark suppression factor. Since the gluon is supposed to fragment more softly than a quark jet, the fragmentation fuction may be chosen independently. Another common option is to split the g jet into a pair of parallel q and $\overline{q}$ ones, sharing the energy, e.g. according to the Altarelli-Parisi splitting function in eq. (20). The fragmentation function could still be chosen independently, if so desired. Further, in either case the fragmentation p_T could be chosen to have a different mean.

Conservation and Lorentz Covariance Issues:
The concept of IF inevitably leads to the total flavour, energy and momentum not being exactly conserved in the fragmentation process proper. At the end of the generation, special algorithms are therefore used to patch this up. The choice of approach has major consequences, e.g. for event shapes and α_S determinations.

Little attention is usually given to flavour conservation. Typically, that aspect is solved by reassigning the flavour content of centrally produced particles, without changing their three-momenta, so that the net number of each flavour is vanishing.

Several different schemes for energy and momentum conservation have been devised. One [48] is to conserve transverse momentum locally within each jet, so that the final momentum vector of a jet is always parallel with that of the corresponding parton. Then longitudinal momenta may be rescaled separately for particles within each jet, such that the ratio of rescaled jet momentum to initial parton momentum is the same in all jets. Since the initial partons had net vanishing three-momentum, so do now the hadrons. The rescaling factors may be chosen such that also energy comes out right.

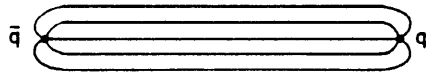

Figure 7: A uniform colour flux tube stretched between a q and $\overline{q}$ endpoint — one possibility of visualizing the linear confinement property.

Another common approach [49] is the one used in EUROJET, in which a generated event is boosted to the frame where the total hadronic momentum is vanishing. After that, energy conservation can be obtained by rescaling all particle three-momenta by a common factor.

A serious conceptual weakness of the IF framework is the issue of Lorentz invariance. The outcome of the fragmentation procedure depends on the coordinate frame chosen, a problem circumvented by requiring fragmentation always to be carried out in the CM frame. This is a consistent procedure for two-jet events, but only a technical trick for multijets.

It should be noted, however, that a Lorentz covariant generalization of the independent fragmentation model exists, in which separate 'gluon-type' and 'quark-type' strings are used, the Montvay scheme [51]. For a three-jet event, the three string pieces are joined at a junction. The motion of this junction is given by the composant of the string tensions acting on it. In particular, it is always possible to boost an event to a frame where this junction is at rest. In this frame, much of the standard naive IF picture holds for the fragmentation of the three jets; additionally, a correct treatment would automatically give flavour, energy and momentum conservation. Unfortunately, while the scheme is perfectly valid also with several gluon jets present, it is then very complicated to write an event generator based on it, and nobody has ever done so.

A second conceptual weakness of IF is the issue of collinear divergences. In a parton shower picture, where a quark or gluon is expected to branch into several reasonably collimated partons, the independent fragmentation of one single parton or of a bunch of collinear ones gives quite different outcomes, e.g. with a much larger hadron multiplicity in the latter case. It is conceivable that a different set of fragmentation functions could be constructed in the shower case in order to circumvent this problem (local parton-hadron duality would correspond to having $f(z) = \delta(z-1)$), but so far no explicit model has been constructed along these lines.

2.3.2 String Fragmentation

The first example of a string fragmentation (SF) scheme was given by Artru and Mennessier [52]. With the elaborate string model developed by the Lund group in the years around 1980 [53], SF became more or less synonymous with the Lund model. This is perhaps less true than it used to be: string ideas are nowadays often used in cluster fragmentation, while the firestring model of EPOS represents another related concept.

The string concept:
In QCD, a linear confinement is expected at large distances (neglecting the effect of string breaks). This provides the starting point for the string model, most easily il-

lustrated for the production of a back-to-back $q\overline{q}$ jet pair. As the partons move apart, the physical picture is that of a colour flux tube (or maybe colour vortex line) being stretched between the q and the $\overline{q}$, Fig. 7. The transverse dimensions of the tube are of typical hadronic sizes, roughly 1 fm. If the tube is assumed to be uniform along its length, this automatically leads to a confinement picture with a linearly rising potential. In order to obtain a Lorentz covariant and causal description of the energy flow due to this linear confinement, the most straightforward way is to use the dynamics of the massless relativistic string with no transverse degrees of freedom [54]. The mathematical, one-dimensional string can be thought of as parametrizing the position of the axis of a cylindrically symmetric flux tube. From hadron mass spectroscopy the string constant, i.e. the amount of energy per unit length, is deduced to be $\kappa \approx 1$ GeV/fm. The expression 'massless' relativistic string is somewhat of a misnomer: κ effectively corresponds to a 'mass density' along the string.

If several partons are moving apart from a common origin, the details of the string drawing become more complicated. For a $q\overline{q}g$ event, a string is stretched from the q end via the g to the $\overline{q}$ end, i.e. the gluon is a kink on the string, carrying energy and momentum. As a consequence, the gluon has two string pieces attached, and the ratio of gluon/quark string force is two, a number which can be compared with the ratio of colour charge Casimir operators, $N_C/C_F = 2/(1-1/N_C^2) = 9/4$. In this, as in other respects, the string model can be viewed as a variant of QCD where the number of colours N_C is not 3 but infinite. Note that the factor 2 above does not depend on the kinematical configuration: a smaller opening angle between two partons corresponds to a smaller string length drawn out per unit time, but also to an increased transverse velocity of the string piece, which gives an exactly compensating boost factor in the energy density per unit string length.

In an event with several gluons, these will still appear as kinks on the string between the q and $\overline{q}$ ends. With several gluons present, the full string evolution may become rather complicated, in particular when two partons connected with a string segment have a small invariant mass. The net effect, however, is that two such nearby partons together drag out a string very much like what would have been dragged out by one single parton with the summed momentum. A soft gluon, on the other hand, does not affect the string evolution significantly. These properties of the string motion are the reasons why the string fragmentation scheme is 'infrared safe' with respect to soft or collinear gluon emission.

The fragmentation process:
Let us now turn to the fragmentation process, and for simplicity consider a $q\overline{q}$ two-jet event. As the q and $\overline{q}$ move apart, the potential energy stored in the string increases, and the string may break by the production of a new $q'\overline{q}'$ pair, so that the system splits into two colour singlet systems $q\overline{q}'$ and $q'\overline{q}$. If the invariant mass of either of these string pieces is large enough, further breaks may occur.

- In the Lund string model, the string breakup process is assumed to proceed until only on-mass-shell hadrons remain, each hadron corresponding to a small piece of string.
- In CALTECH-II, on the contrary, the breaking is stopped at a string piece mass of a few GeV, and these pieces are identified with clusters.

In order to generate the quark-antiquark pairs $q'\overline{q}'$ which lead to string breakups, the Lund model invokes the idea of quantum mechanical tunnelling. In terms of the transverse mass m_T of the q', the tunnelling probability (i.e. the probability that the $q'\overline{q}'$ will appear) is given by

$$\exp\left(-\frac{\pi m_T^2}{\kappa}\right) = \exp\left(-\frac{\pi m^2}{\kappa}\right)\exp\left(-\frac{\pi p_T^2}{\kappa}\right). \tag{29}$$

The factorization of the transverse momentum and the mass terms leads to a flavour-independent Gaussian spectrum for the p_T of $q'\overline{q}'$ pairs. Since the string is assumed to have no transverse excitations, this p_T is locally compensated between the quark and the antiquark of the pair, just as in IF. In a perturbative QCD framework, a hard scattering is associated with gluon radiation, and further contributions to what is naively called fragmentation p_T comes from unresolved radiation. This is used as an explanation why the experimental $\langle p_T \rangle$ is somewhat higher than obtained with the formula above.

The formula also implies a suppression of heavy quark production $u : d : s : c \approx 1 : 1 : 0.3 : 10^{-11}$. Charm and heavier quarks are hence not expected to be produced in the soft fragmentation. Since the predicted flavour suppressions are in terms of quark masses, which are notoriously difficult to assign (should it be current algebra, or constituent, or maybe something in between?), the suppression of $s\overline{s}$ production is left as a free parameter in the program. At least qualitatively, the experimental value agrees with theoretical prejudice. As for IF, it is necessary to invoke an algorithm to choose between pseudoscalar and vector mesons. Here the string model is not particularly predictive. Qualitatively one expects a $1:3$ ratio, from counting the number of spin states, multiplied by some wave function normalization factor, which should disfavour heavier states. Again, the relative composition is therefore left as a free parameter in Monte Carlo implementations.

A tunnelling mechanism can also be used to explain the production of baryons. This is still a poorly understood area. One possible approach is described in the JETSET section.

In contrast, the CALTECH-II model possesses no mechanism to generate transverse momenta for the q' and $\overline{q}'$ produced in string breakups. All p_T comes from shower evolution and cluster decays (see below); all flavour suppression comes from phase-space effects.

Fragmentation functions:
In general, the different string breaks are causally disconnected. This means that it is possible to describe the breaks in any suitable order, e.g. from a quark end inwards. What was just an ansatz in the independent fragmentation approach is thereby given some underlying physical justification.

A fragmentation process described in terms of starting at the q end of the system and fragmenting towards the $\overline{q}$ end should be equivalent to doing it the other way around. This 'left-right' symmetry constrains the allowed shape of fragmentation functions $f(z)$, where (for two-jets) z is once again the fraction of $E + p_L$ along the jet axis chosen. With some simplifying assumptions, the left-right symmetric fragmentation function takes the form

$$f(z) \propto z^{-1}(1-z)^a \exp(-bm_T^2/z), \tag{30}$$

with the two free parameters a and b.

If the objects produced during string breakup are not on-shell hadrons, as is the case for CALTECH-II, the logical recipe is to have a constant probability for the string to break per unit of (invariant) area swept out by the string. This gives an exponential area decay law, and also ensures left-right symmetry. The specification of adjacent breakups in the two-dimensional sheet of the string can be translated into a simultaneous probability distribution in the mass-squared and the z of the produced cluster:

$$\frac{d\mathcal{P}}{dm^2 dz} \propto z^{-1} \exp(-bm^2/z), \tag{31}$$

with b a free parameter proportional to the probability of breakup per invariant unit of string area. The mass-spectrum can be obtained by integrating out z, and is proportional to $E_1(bm^2)$, i.e. it posseses a logarithmic divergence at small masses (in programs often removed by a minimum cluster mass requirement). With the m^2 given, the fragmentation function can just be read off eq. (31), i.e. it is exactly the same as was obtained in the discrete mass case, eq. (30), except that $a = 0$. The similarity in the final result is all the more surprising, since a discrete mass spectrum implies that string breaks are only allowed along one-dimensional hyperbolae, rather than inside two-dimensional areas.

The UCLA model for flavour composition:
Perhaps the biggest weakness of the Lund string picture is its inability to predict particle composition in terms of experimental observables, like hadron masses. Instead it predicts particle composition in terms of uncertain quantities like quark or diquark masses. The UCLA extension of the Lund model is an attempt to solve that problem, starting with the Lund symmetric fragmentation function in eq. (30). The standard Lund interpretation is to view this as the probability to pick a z value, once the choice of hadron and p_T has already been made; i.e. to fix the normalization to unit probability separately for each hadron. In the UCLA picture [55], the normalization is assumed to be common, which means that the relative probability to produce a hadron with given mass is proportional to the integral over allowed z and p_T values:

$$\mathcal{P}(m^2) \propto \int_0^1 dz \int dp_T^2 \frac{1}{z}(1-z)^a \exp\left(-\frac{b(m^2+p_T^2)}{z}\right). \tag{32}$$

Besides the integral above, the only additional weight factors are trivial Clebsch-Gordan coefficients and spin counting factors. In particular, there is no quark level suppression of s quarks, c quarks (!), or diquarks.

With this framework, it is indeed possible to get a surprisingly good description of the flavour composition in e^+e^- events, only using the standard a and b parameters already present in the fragmentation function, plus a fudge factor to enhance baryon production in general. (In recent modelling, this fudge factor has been replaced by an extended baryon production scenario, more based on first principles.) This does not mean that agreement with data is perfect, so the picture may still be too simplistic.

The UCLA model in principle also predicts the transverse momentum spectrum for hadrons, but without giving a recipe for how this transverse momentum should be globally conserved. To date, no consistent scheme is known for solving this problem, so the standard Lund recipe for generating p_T on the parton level is probably to be preferred.

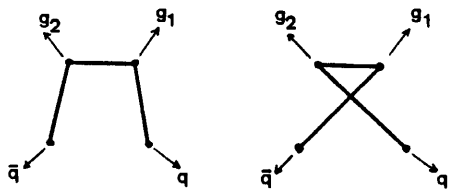

Figure 8: The two different ways of stretching a string in a $q\bar{q}gg$ event.

Interfacing to matrix elements and parton showers:
For four-jet events (or events with more than four jets) which are generated using matrix elements, there are several possible topologies for the ordering of partons along the string. This is illustrated in Fig. 8 for $q\bar{q}gg$ events. A knowledge of quark and gluon colours, obtained by perturbation theory, would uniquely specify the stretching of the string [56], so long as the two gluons do not have the same colour. The probability for the latter is down in magnitude by a factor $1/N_C^2$, where $N_C = 3$ is the number of colours. Perturbative QCD gives no answer for how to handle these situations, but recipes have to be included in event generators.

In leading log shower programs, where only $1 \to 2$ branchings are included, the rules for colour flow in branchings are well-defined. If one goes beyond this approximation, as in NLLjet, the same problems will arise as in the four-jet matrix elements, and similar solutions have to be found.

Whichever scenario is considered, matrix elements or showers, there is a tacit assumption that soft gluon exchanges between partons will not mess up the original colour assignment. Both theoretical and experimental arguments can be raised as to why this might be a fair approximation, but eventually this assumption is something which needs more rigorous testing [57].

2.3.3 Cluster Fragmentation

While the Artru-Mennessier model [52] contained clusters, dedicated cluster models were first developed at CALTECH [41]. Today, cluster models are found in HERWIG and in CALTECH-II. The main difference between cluster fragmentation (CF) schemes is the extent to which string fragmentation ideas are incorporated.

- In one extreme, a parton shower picture is used to produce a partonic configuration. At the end of the shower evolution, remaining gluons are forcibly split into $q\bar{q}$ pairs. With colour explicitly kept track of, the quark of one splitting may be combined with the antiquark from an adjacent one to form a colourless cluster, Fig. 9. These clusters subsequently decay into the final hadrons. This is, more or less, the HERWIG strategy.
- In the other extreme, parton showers or matrix elements may be used to generate a partonic configuration, with string stretched in between the partons, as described in the SF section. These strings then fragment into clusters, which again decay into the final hadrons. Fig. 10 illustrates this approach, which is found in CALTECH-II.

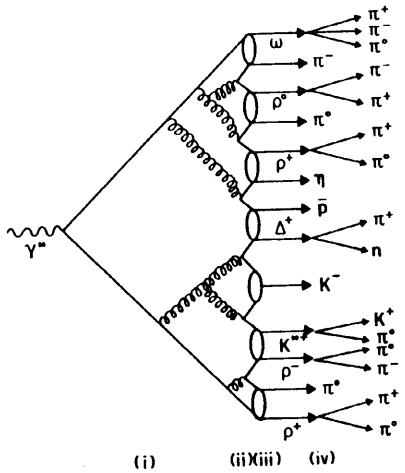

Figure 9: One cluster fragmentation scenario: (i) shower evolution, (ii) forced $g \to q\bar{q}$ branchings, (iii) cluster formation, and (iv) cluster decay. Occasionally, a cluster is associated with a single particle.

The cluster concept:
The concept of cluster fragmentation offers the great promise of a simple, local and universal description of hadronization. Gone are the long, ordered fragmentation chains present both in SF and IF. In their place appear simple clusters, which are assumed to be the basic units from which the hadrons are produced. A cluster is ideally only characterized by its total mass and total flavour content, i.e. unlike a string it does not possess an internal structure. If the shower evolution and/or string breaks are chosen such that most clusters have a mass of a few GeV, the cluster mass spectrum may be thought of as a superposition of fairly broad (i.e. short-lived) resonances.

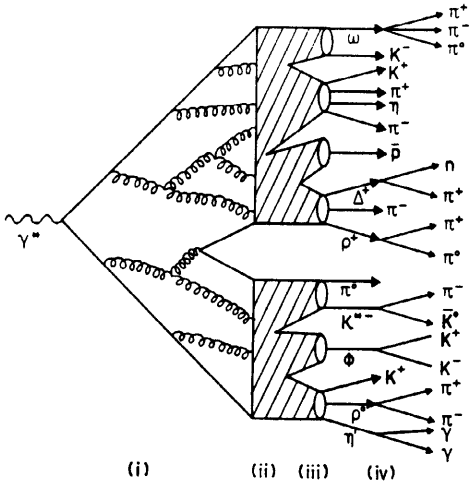

Figure 10: Another cluster fragmentation scenario: (i) shower evolution, (ii) string evolution and breakup into (iii) clusters, (iv) cluster decay into a varying number of particles.

Phase-space aspects may then be expected to dominate the decay properties. This applies both for the selection of decay channels, and for the kinematics of the decay. Thus a decay is assumed to be isotropic in the rest frame of the cluster. This gives a compact description with few parameters. In particular, the separate longitudinal and transverse momentum fragmentation descriptions in SF and IF are here replaced by a unified framework, wherein parton showers and cluster phase-space decays give the full momentum distribution. As we shall see in the following paragraphs, there are some complications, however.

Ideally, parton shower evolution should in itself give a cluster mass spectrum strongly damped at masses above a few GeV, so that two-body decays of clusters would give a sufficient description. This was the hope in the early days of CF, a hope founded on the concept of 'preconfinement' [58]. Unfortunately, despite the preconfinement property, there is no known way of avoiding a rather large spread of cluster masses. In programs, it is therefore necessary to introduce the possibility for a high-mass cluster to produce more than two hadrons. This is typically done by allowing branchings cluster $\to$ cluster + hadron or cluster $\to$ cluster + cluster. Usually, these decays are not performed isotropically, but as if the decaying cluster were a 'ministring' with a well-defined longitudinal direction.

Flavour generation:
Flavours are generated at several different stages. First, at the branchings $g \to q\bar{q}$, alternatively at the string breaks, i.e. when the clusters are formed. In CALTECH-II the relative probabilities appear as explicit parameters, while they are given by the parton mass assignments in HERWIG. In the original Webber-Marchesini model, only quark-antiquark pairs (or gluons) were created at branchings. This leads to some problems in the description of baryon production. The present model therefore also includes, as an option, the possibility that a gluon may branch into a diquark-antidiquark pair. This process is turned on below some scale in the shower evolution, with an arbitrary strength relative to ordinary $q\bar{q}$ production.

A second stage of flavour production occurs when larger clusters decay into smaller ones. Typically, this means that a cluster $q_1\bar{q}_2$ breaks, by the production of an intermediate $q_3\bar{q}_3$ pair, into clusters $q_1\bar{q}_3$ and $q_3\bar{q}_2$. One of the two may, but need not, be directly associated with a hadron. In general, the symbol q may here represent either a quark or an antidiquark, but in HERWIG production of a new diquark-antidiquark pair is forbidden at this stage. The third stage of flavour production is when a cluster decays into two hadrons. The flavour flow is as above, i.e. a new $q_3\bar{q}_3$ pair splits the old cluster in the middle. Here quark, diquark, and even charm production is allowed in HERWIG, with relative probability dictated by phase-space alone.

The phase-space assumption means that each allowed cluster decay channel is assigned a weight proportional to the density of states, i.e. equal to $(2s_1 + 1)(2s_2 + 1)(2p^*/m)$. Here s_1 and s_2 are the spins of the two hadrons produced, and p^* the common momentum of the products in the rest frame of the decaying cluster (with mass m). The weight gives the relative probability of the choice being retained; in case of rejection a new q_3 flavour is selected and the procedure repeated.

It should be noted that the new q_3 quark flavour is not associated with any dynamical properties, such as a mass or, for diquarks, a total spin. It is only the properties of the final, 'observable' particles that can influence the relative production rate. Further, the

'fragmentation' transverse momentum is determined by the average energy release in cluster decay, and in subsequent resonance decays, as opposed to the extra parameter needed in the Lund model.

Heavy and light clusters:
In the decay of a large cluster into two, the kinematics is usually handled anisotropically, along the 'string' direction. The same kind of phase-space weight may still be used, provided clusters are assigned suitable spins and a cluster mass spectrum weight is folded in. This is what is done in CALTECH-II, while the HERWIG breakup rule is a simpler phenomenological recipe.

If all clusters are to decay into at least two particles, the probability of producing a single particle carrying a large fraction of the total jet energy is severely underestimated. Therefore cluster programs usually contain a mechanism, so that a sufficiently light cluster is assumed to collapse into a single particle. Four-momentum is shuffled to or from nearby clusters, so as to achieve overall energy and momentum conservation.

2.3.4 Other Fragmentation Approaches

As should be amply clear from the discussions above, the study of fragmentation is not a closed chapter, with one simple framework that does it all. In addition, none of the models discussed are able to describe some experimental features. This means that a broad spectrum of alternatives should be pursued, a few of which appear in the program descriptions. Many of these are toy models, useful for illustrating principles, but not for the detailed description of experimental data, although at least one, EPOS, is a fully-fledged program in its own right. Other models are not embodied in computer programs, but may still provide interesting insights. In the following paragraphs, we introduce a few interesting approaches, without any ambition of completeness.

Local Parton-Hadron Duality:
One of the most powerful approaches, and yet conceptually a simple one, is to assume that properties on the parton level are intimately related to the corresponding hadronic properties, to the extent that (some) quantities may be studied without making any reference at all to a fragmentation process. This idea of a local parton-hadron duality has been developed in particular by the Leningrad group [35].

In the 'hard-line' interpretation, a one-to-one correspondence is assumed between partons and hadrons, event by event, given only that the partonic cut-off parameter is suitably chosen. In the 'soft-line' interpretation, which most people seem to favour, there is no absolute correspondence event by event, but only in the average behaviour, and in the fluctuations around that average. This framework can e.g. be used to predict correlations of the '$E^\alpha M^\beta C$' kind, which means that one should fix α jet directions from the energy flow, and then consider M-particle multiplicity correlations, internally and with respect to the jet directions (thus the ordinary string studies would be E^3MC, i.e. specifying three jet directions and then considering the inclusive particle flow) [35]. The procedure of normalizing to a given jet energy configuration will ensure that results are finite and well-behaved.

EPOS:
EPOS is not based on QCD, or at least not on QCD as we know it. Rather, its underlying

assumption is that there are no dynamical gluons. Only $q\bar{q}$ (or qqq etc.) bound states are physical entities. Lower-mass states correspond to the ordinary hadrons, while higher-mass states are called fire-strings. The 'string' suffix is related to the fact that high angular momenta are generated for these states, which give them an anisotropic structure in their decays, just as for ordinary strings. In an e^+e^- annihilation event, a fire-string is produced at the primary γ/Z^0 vertex. This initial fire-string may break into two fire-strings, into a hadron plus a fire-string, or into two hadrons. By successive branchings towards smaller masses, a purely hadronic final state is obtained. The probability that fire-strings are 'tilted' with respect to their mother fire-string direction is what produces a non-two-jet structure, without ever resorting to the ansatz of gluon emission.

Statistical models:
The model of Ochs (DPSJET) is a purely statistical model, just based on the equipartitioning of energy in the fragmentation process. No QCD whatsoever is used, and yet multijet events do appear, and an effective α_S value that is a prediction of the model. Since only pions can be created, and since there are some problems describing multiplicity, DPSJET cannot directly be compared with data. Nonetheless, some of its properties could indeed be studied experimentally.

An intermittency model:
Another toy model is the program being developed by the Cracow group. The model is only implemented in one space dimension, which makes it unsuitable for comparisons with data, but the issue considered is the more important: to what extent could final state interactions mess up the fragmentation models described previously? To study this question, a string type fragmentation picture is invoked. The fragmenting string pieces are allowed to reinteract, should their paths cross. The central issue is whether this kind of reinteraction could reproduce intermittency phenomena.

A baryon production model:
One of the least well understood aspects of particle production is that of baryon formation. Most models treat this in the same spirit as meson formation, but maybe the underlying physics is quite different. In the model of Ellis and Kowalski, baryons are produced at topological defects [59]. These defects appear when chiral symmetry is broken during the fragmentation phase, and adjacent domains (corresponding to the final mesons) have a mismatch in chiral directions. In this model, the baryon multiplicity should increase faster with CM energy than in conventional approaches, which should produce visible differences at LEP.

2.4 Particles and Their Decays

Particle content:
All the main QCD Monte Carlos include the lowest-lying pseudoscalar and vector mesons, and most contain the lowest-lying spin 1/2 ('octet') and 3/2 ('decuplet') baryon states. Some programs also include mesons with orbital angular momentum $L = 1$, i.e. the scalar, pseudovector (2 of them) and tensor multiplets, as well as a few radial excitations, type Ψ'. Excited baryon states are usually not included.

Normally all particles in the full three-generation standard model are included; some-

times a fourth generation is also included. If any of these particles are found at LEP, the expected lifetimes are long enough for hadrons to actually be formed. The full menagerie of possible mesons and baryons is extensive, and are usually treated *en masse* for the heavier quark flavours, according to schematic rules. This applies both for masses and decays.

Masses:
Known particle masses are taken from the Review of Particle Properties [14], while unknown masses are calculated according to simple mass formulae.

Many particles are short-lived resonances, for which the mass value is not well-defined, e.g. ρ, K^*, or Δ. Some programs do include a Breit-Wigner type broadening for these resonances, but none includes the full machinery of correct threshold factors. The reason is that a correct treatment requires a coupled description of production and decay, while most programs are based on a sequential description, where the particle is first produced with some given mass, and only at a later stage decays according to fixed branching ratios. An (extreme) example to illustrate why sequential treatment is not correct in general: the scalar meson f_0 has a nominal mass below the K^+K^- threshold, and therefore predominantly decays to $\pi^+\pi^-$ and $\pi^0\pi^0$. The resonance is broad enough, however, for the more massive specimens to have a chance to decay to K^+K^- or even $K^0\overline{K}^0$. The resonance shape is therefore not of the simple Breit-Wigner type, and branching ratios vary wildly with mass.

In general, existing Monte Carlo programs should therefore only be used as a first guideline for what to expect in the search for a broad resonance. The detailed modelling is left to the experimenter, for the particular channel under study.

Decays:
It is a general misconception that the treatment of decays is fairly routine, similar between programs. True, the differences are smaller in this area than they are in the perturbative QCD or fragmentation areas, but that does not make them negligible, as the following comments illustrate.

Branching ratios are taken, where measured, from the Review of Particle Properties [14]. For light particles, experimental knowledge is fairly complete, and programs do not differ by much. Particles are assumed to decay isotropically, according to phase-space. Some programs include additional matrix elements, which multiply the ordinary phase-space weights. This may include π^0 and η Dalitz decays to γe^+e^-, ω and ϕ decays to $\pi^+\pi^-\pi^0$, and vector decays in the sequence pseudoscalar $\rightarrow$ pseudoscalar + vector $\rightarrow$ 3 pseudoscalars.

For heavier particles, the situation is more complicated. Experimental knowledge is no longer complete. This means that each program builder has to make educated guesses. One example is the τ puzzle, i.e. that experimentally the sum of branching ratios for exclusive one-prong decays is lower than the inclusive one-prong branching ratio. Many brute force solutions are possible here, and users should not expect any two programs to agree.

For D^0 and D^+, established branching ratios sum up to roughly 80% [60]. With symmetry arguments and a few educated guesses, reasonably complete sets of decay channels and branching ratios may be obtained, but certainly with differences from program to program. These differences are even larger for D_s and for charmed baryons,

where few branching ratios are measured. It is still possible to enumerate a closed set of decay channels. Some programs take this approach, while others make do with simplified statistical models, whereby a multiplicity is picked according to some distribution, and the requested number of particles is constructed from an original flavour content, given by the weak decay process, plus a number of additional $q'\overline{q}'$ pairs produced at random. For bottom decays, the fraction of measured branching ratios is vanishingly small, while the set of possible decay channels is very large. Therefore all programs resort to statistical models, and the differences can be significant.

For weak decays, in particular semileptonic ones, it would not do to distribute particles entirely according to phase-space. On the parton level, the matrix elements are known. For a charge +2/3 quark decay $Q \rightarrow q + l^+ + \nu$,

$$|M|^2 \propto (p_Q p_l)(p_\nu p_q), \tag{33}$$

while for a charge -1/3 quark decay $Q \rightarrow q + l^- + \overline{\nu}$,

$$|M|^2 \propto (p_Q p_\nu)(p_l p_q). \tag{34}$$

The matrix elements above generalize to hadronic quark decay modes, and to τ decays, with simple substitutions so as to preserve the weak isospin nature.

In the end, it is again necessary to translate partons, distributed according to the matrix elements above, into hadrons. Often the decay product (q above) combines with the spectator quark to form a single hadron (in B decays, the modes B $\rightarrow$ (D, D^*) l^- ν_l are known almost to saturate the semileptonic branching ratio). In hadronic B decays, it is therefore usual to assume the production of a single D or D^*, plus a phase-space model for the decays of the quarks of the virtual W.

Most programs do not worry about particle polarization. Such polarization could be present for a number of reasons: it could be inherited from the initial $q\overline{q}$ pair produced at the γ/Z^0 vertex (particularly relevant for heavy flavours), it could be produced in the fragmentation process (as we know is the case with Λ baryons and other hyperons), or it could arise in weak decays (with τ being the most interesting case). EURODEC and TIPTOP do contain helicity information, which is used, e.g., in τ decays, and thus are more sophisticated with regard to their treatment of hadron decay phenomenology than are most models.

References

[1] Report of the Electroweak theory working groups

[2] Report of the Electroweak Event Generators working group

[3] Report of the QCD working group

[4] G. d'Agostini, W. de Boer, G. Grindhammer, DESY 89-057 (1989)

[5] M. Dine, J. Sapirstein, Phys. Rev. Lett. **43** (1979) 668;
K.G. Chetyrkin *et. al.*, Phys. Lett. **B85** (1979) 277;
W. Celmaster, R.J. Gonsalves, Phys. Rev. Lett. **44** (1980) 560

[6] S.G. Gorishny, A.L. Kataev, S.A. Larin, Phys. Lett. **B212** (1988) 238

[7] Report of the Heavy Flavour working group

[8] H.A. Olsen, P. Osland, I. Øverbø, Nucl. Phys. **B171** (1980) 209;
J. Jersák, E. Laermann, P.M. Zerwas, Phys. Rev. **D25** (1982) 1218

[9] A. Ali, J.G. Körner, Z. Kunszt, E. Pietarinen, G. Kramer, G. Schierholz, J. Willrodt, Nucl. Phys. **B167** (1980) 454

[10] J.G. Körner, G.A. Schuler, Z. Physik **C26** (1985) 559, DESY 87-124 (1987)

[11] E. Laermann, T.F. Walsh, I. Schmitt, P.M. Zerwas, Nucl. Phys. **B207** (1982) 205

[12] S. Jadach, Z. Wąs, Phys. Lett. **B219** (1989) 103

[13] J. Ellis, M.K. Gaillard, G.G. Ross, Nucl. Phys. **B111** (1976) 253

[14] Particle Data Group, G.P. Yost et al., Phys. Lett. **B204** (1988) 1

[15] P.M. Stevenson, Phys. Rev. **D23** (1981) 2916

[16] K.J.F. Gaemers, J.A.M. Vermaseren, Z. Physik **C7** (1980) 81

[17] R.K. Ellis, D.A. Ross, A.E. Terrano, Nucl. Phys. **B178** (1981) 421

[18] D. Danckaert, P. De Causmaecker, R. Gastmans, W. Troost, T.T. Wu, Phys. Lett. **B114** (1982) 203

[19] B.A. Kniehl, J.H. Kühn, Phys. Lett. **B224** (1989) 229

[20] J.A.M. Vermaseren, K.J.F. Gaemers, S.J. Oldham, Nucl. Phys. **B187** (1981) 301

[21] K. Fabricius, G. Kramer, G. Schierholz, I. Schmitt, Z. Physik **C11** (1982) 315

[22] Z. Kunszt, Phys. Lett. **B99** (1981) 429, **B107** (1981) 123;
T.D. Gottschalk, Phys. Lett. **B109** (1982) 331;
R.-y. Zhu, MIT Ph.D. thesis (1983);
A. Ali, F. Barreiro, Nucl. Phys. **B236** (1984) 269;
F. Gutbrod, G. Kramer, G. Schierholz, Z. Physik **C21** (1984) 235;
T.D. Gottschalk, M.P. Shatz, Phys. Lett. **B150** (1985) 451, CALT-68-1172 (1984);
G. Kramer, B. Lampe, Fortschr. Phys. **37** (1989) 161;
F. Gutbrod, G. Kramer, G. Rudolph, G. Schierholz, Z. Physik **C35** (1987) 543

[23] JADE Collaboration, S. Bethke *et. al.*, Phys. Lett. **B213** (1988) 235;
TASSO Collaboration, W. Braunschweig *et. al.*, Phys. Lett. **214B** (1988) 286

[24] W. de Boer, DESY 89-067, to be published in Proc. of Ringberg Workshop, Munich, ed. J.H. Kühn (1989);
W. de Boer, SLAC-Pub 4482, Proc. of the Xth WARSAW Symposium on Elementary Particle Physics, Kazimierz, Poland, ed. Z. Ajduk (1987), p. 503

[25] G. Kramer, B. Lampe, Z. Physik **C39** (1988) 101

[26] N. Magnussen, DESY F22-89-01, Ph.D. Thesis (1989)

[27] CELLO Collaboration, H. J. Behrend *et. al.*, contribution to the XXIV International Conference on High Energy Physics, Munich, 1988

[28] S. Bethke, LBL-26958 (1989)

[29] K. Hagiwara, D. Zeppenfeld, Nucl. Phys. **B313** (1989) 560;
N.K. Falck, D. Graudenz, G. Kramer, Phys. Lett. **B220** (1989) 299, DESY 89-046 (1989)

[30] F.A. Berends, W.T. Giele, H. Kuijf, Leiden preprint 98-0055 (1988)

[31] C. J. Maxwell, Durham preprint DTP/89/08 (1989)

[32] L.A. Gribov, E.M. Levin, M.G. Ryskin, Phys. Rep. **100** (1983) 1;
A. Bassetto, M. Ciafaloni, G. Marchesini, Phys. Rep. **100** (1983) 201;
B.R. Webber, Ann. Rev. Nucl. Part. Sci. **36** (1986) 253

[33] G. Altarelli, G. Parisi, Nucl. Phys. **B126** (1977) 298

[34] A.H. Mueller, Phys. Lett. **B104** (1981) 161;
B.I. Ermolaev, V.S. Fadin, JETP Lett. **33** (1981) 269

[35] Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Sov. J. Nucl. Phys. (Yad. Fiz.) **47** (1988) 1384, DESY 88-093 (1988)

[36] JADE Collaboration, W. Bartel *et. al.*, Phys. Lett. **B101** (1981) 129, **B134** (1984) 275, Z. Physik **C21** (1983) 37;
TPC/2γ Collaboration, H. Aihara *et. al.*, Z. Physik **C28** (1985) 31;
TASSO Collaboration, M. Althoff *et. al.*, Z. Physik **C29** (1985) 29

[37] B. Andersson, G. Gustafson, T. Sjöstrand, Phys. Lett. **B94** (1980) 211

[38] A. Bassetto, M. Ciafaloni, G. Marchesini, A.H. Mueller, Nucl. Phys. **B207** (1982) 189;
J.B. Gaffney, A.H. Mueller, Nucl. Phys. **B250** (1985) 109

[39] K. Kajantie, E. Pietarinen, Phys. Lett. **B93** (1980) 269

[40] C.-H. Lai, J.L. Petersen, T.F. Walsh, Nucl. Phys. **B173** (1980) 244

[41] G.C. Fox, S. Wolfram, Nucl. Phys. **B168** (1980) 285;
R.D. Field, S. Wolfram, Nucl. Phys. **B213** (1983) 65

[42] J.C. Collins, Nucl. Phys. **B304** (1988) 794;
I.G. Knowles, Nucl. Phys. **B310** (1988) 571

[43] J. Kalinowski, K. Konishi, T.R. Taylor, Nucl. Phys. **B181** (1981) 221;
J.F. Gunion, J. Kalinowski, L. Szymanowski, Phys. Rev. **D32** (1985) 2303

[44] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Phys. Lett. **B165** (1985) 147, Leningrad preprint 1051 (1985)

[45] T. Sjöstrand, Int. J. Mod. Phys. **A3** (1988) 751

[46] A. Krzywicki, B. Petersson, Phys. Rev. **D6** (1972) 924;
J. Finkelstein, R.D. Peccei, Phys. Rev. **D6** (1972) 2606;
F. Niedermayer, Nucl. Phys. **B79** (1974) 355;
A. Casher, J. Kogut, L. Susskind, Phys. Rev. **D10** (1974) 732

[47] R.D. Field, R.P. Feynman, Nucl. Phys. **B136** (1978) 1

[48] P. Hoyer, P. Osland, H.G. Sander, T.F. Walsh, P.M. Zerwas, Nucl. Phys. **B161** (1979) 349

[49] A. Ali, J.G. Körner, G. Kramer, J. Willrodt, Nucl. Phys. **B168** (1980) 409;
A. Ali, E. Pietarinen, G. Kramer, J. Willrodt, Phys. Lett. **B93** (1980) 155

[50] C. Peterson, D. Schlatter, I. Schmitt, P. Zerwas, Phys. Rev. **D27** (1983) 105

[51] I. Montvay, Phys. Lett. **B84** (1979) 331

[52] X. Artru, G. Mennessier, Nucl. Phys. **B70** (1974) 93

[53] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rep. **97** (1983) 31

[54] X. Artru, Phys. Rep. **97** (1983) 147

[55] C.D. Buchanan, S.-B. Chun, Phys. Rev. Lett. **59** (1987) 1997, UCLA-HEP-89-003 (1989)

[56] G. Gustafson, Z. Physik **C15** (1982) 155

[57] G. Gustafson, U. Pettersson, P. Zerwas, Phys. Lett. **B209** (1988) 90

[58] D. Amati, G. Veneziano, Phys. Lett. **B83** (1979) 87;
G. Marchesini, L. Trentadue, G. Veneziano, Nucl. Phys. **B181** (1981) 335

[59] J. Ellis, H. Kowalski, Phys. Lett. **B214** (1988) 161, CERN-TH.5316/89 and NEVIS R# 1405 (1989)

[60] D. G. Hitlin, Nucl. Phys. **B3** (proc. supp.) (1988) 179

3 Matrix Element Programs

In section 2.2.1, we have given an introduction to the matrix element approach. On the Monte Carlo side, no standard implementation of these matrix elements has emerged: among the early programs, Hoyer *et. al.* [1] and JETSET only contained first order QCD, and Ali *et. al.* [2], while containing four-jets, still did not include second order corrections to the three-jet rate. In later versions, JETSET did contain full second order matrix elements, but according to the GKS scheme [3], which is nowadays generally recognized to miss some non-negligible terms. Other programs, like COJETS or the Marchesini-Webber or Gottschalk ones, were intended for shower evolution from the start, and never contained matrix element options.

Instead, different generators have been developed by experimentalists. These have mainly been used inside the original collaborations, have often not been publicly available, and not always well documented. In this section we will present some of these algorithms. More may exist that we are not aware of.

Essentially all generators make use of some simplifications to the full problem.

- When performing the cancellation between soft four-jets and virtual three-jets, the overall angular orientation of events (with respect to the beam axis) is neglected, and only the internal variables of the event are used. Here one in principle knows how to do better [4].
- Calculations of the second order three-jet matrix elements have only been performed for massless quarks. The best one can do here is probably to assume that the effective mass suppression factor (as a function of kinematics) of the first order three-jet matrix elements also carries over to the full second order expression.
- In the $q\bar{q}q'\bar{q}'$ matrix elements, the new pieces only allowed due to the axial nature of the Z^0 [5] are not included.
- If the top quark is heavy, then $b\bar{b}$ production receives non-negligible contributions from loops involving t quarks and W bosons. Programs have no provisions for this. The $b\bar{b}$ rate can be enhanced by brute force, provided that the subsequent relative pattern of gluon emission is assumed to be unchanged.

In practice, none of the limitations above are expected to cause any major uncertainties for LEP applications.

While the approaches used will be discussed in general terms, and compared with each other, our working group have not had the time or the necessary background material to check the correctness of the algorithms used. In view of the very long matrix element expressions that are required (which often have to be typed in by hand), and the delicate cancellation between collapsed four-parton and virtual three-parton events (which is often achieved numerically by Monte Carlo integration), the possibility of errors must always be acknowledged. Quite apart from that, there are indeed ambiguities in the cancellation process, which could well be the full reason why different implementations do give somewhat different results.

3.1 ERT-based Generators (I)

One of the earliest generators was the one implemented by R.-y. Zhu (CALTECH, BITNET ZHU@CITHEX.CALTECH.EDU) [6], which has been extensively used by

Mark-J (and CELLO, see section 3.3). This generator is based on the ERT matrix elements [7], with a Monte Carlo recombination procedure suggested by Kunszt [8] and developed by Ali [9]. It has the merit of giving corrections in a convenient, parametrized form. For practical applications, the main limitation is that the corrections are only given for discrete values of the cut-off parameter(s), y or (ϵ, δ). In the subsequent discussion, we will only refer to y.

The basic approach is the following. Without any loss of generality, the full second order three-jet cross-section can be written in terms of the 'ratio function' $R(X,Y;y)$, defined by

$$\frac{1}{\sigma_0}\frac{d\sigma_3^{tot}}{dX\,dY} = \frac{\alpha_S}{\pi}A_0(X,Y)\left\{1+\frac{\alpha_S}{\pi}R(X,Y;y)\right\}, \qquad (35)$$

where $X = x_1 - x_2 = x_q - x_{\overline{q}}$, $Y = x_3 = x_g$, σ_0 is the lowest order hadronic cross-section, and $A_0(X,Y)$ the standard first order three-jet cross-section, cf. eq. (9). By Monte Carlo integration, the value of $R(X,Y;y)$ is evaluated in bins of (X,Y), and the result parametrized by a simple function $F(X,Y;y)$. Once this function is available, generation of Monte Carlo events is straightforward.

In order to obtain the second order three-jet rate, a small cut $y_0 = 10^{-7}$ was introduced. It was assumed that four-parton events which fail this cut can be (partly) cancelled analytically against the virtual three-jet events, to give a net 'regularized virtual' contribution to the three-jet rate. For a given choice of y cut, in the physical range $y \gg y_0$, an additional 'soft' contribution comes from four-parton events which survive the y_0 cut but fail the y one.

A large sample (9 000 000) of four-parton events was generated inside the y_0 cut region. For events which failed the more stringent y cuts, the parton pair with the smallest invariant mass was recombined into an effective jet, using the $\overline{p}$ recombination scheme. To recapitulate, this means that the individual three-momenta were added, $\overline{p}_{ij} = \overline{p}_i + \overline{p}_j$, the mass of the recombined pair set zero for the calculation of energy, $E_{ij} = |\overline{p}_i + \overline{p}_j|$, and finally all four-momenta were rescaled by a common factor so as to preserve the correct CM frame energy.

In calculating the $\mathcal{O}(\alpha_S^2)$ correction functions, care was taken to maintain the flavour signature of the jets in the recombination process. A quark and a gluon were recombined into a quark with the same flavuor as the original quark, two gluons were recombined to form a gluon, etc. In some cases the three jets of the final state were not in the standard $q\overline{q}g$ configuration. The probability for this to happen corresponded to less than 0.5% of the total cross-section, even for the most stringent cuts used. For these non-$q\overline{q}g$ final states, the assignment of q, $\overline{q}$ and g was done at random.

The sum of 'regularized virtual' (1 000 000 three-jet events were generated, with evaluated second order weights) and 'soft' corrections, normalized to the first order three-jet cross-section, was tabulated in the (X,Y) plane, using bins of size 0.05×0.05. This estimated R function behaviour was then fit with a 12 parameter function F,

$$\begin{aligned} F(X,Y;y) &= p_1 + p_2X^2 + p_3X^4 + (p_4 + p_5X^2)Y + (p_6 + p_7X^2)Y^2 + \\ &\quad (p_8 + p_9X^2)Y^3 + p_{10}/(X^2 - Y^2) + p_{11}/(1-Y) + p_{12}/Y \end{aligned} \qquad (36)$$

The parameters p_i are given in Table 1, for the y values studied. The χ^2 of the F fit to the R table is typically less than 1.1 per degree of freedom.

There are different ways to make use of the parametrization above. The approach of the Zhu algorithm is the following. Three-jet events are generated according to the

Table 1: Coefficients p_1 – p_{12} (first line p_1 – p_4 etc.) of the parametrization in eq. (36), for different y cuts.

$y = 0.01$			
0.1829E+02	0.8956E+02	0.4541E+01	-0.5209E+02
-0.1098E+03	0.2490E+02	0.1163E+02	0.3683E+01
0.1750E+02	0.2440E-02	-0.1362E+01	-0.3537E+00
$y = 0.02$			
0.1142E+02	0.6299E+01	-0.2255E+02	-0.8915E+01
0.5925E+02	-0.5855E+01	-0.3285E+02	-0.1054E+01
-0.1690E+02	0.6489E-02	-0.8156E+00	0.1095E-01
$y = 0.03$			
0.7847E+01	-0.3964E+01	-0.3583E+02	0.1178E+01
0.2939E+02	0.2806E+00	0.4782E+02	-0.1236E+02
-0.5672E+02	0.4054E-01	-0.4365E+00	0.6062E+00
$y = 0.04$			
0.5441E+01	-0.5689E+02	-0.5027E+02	0.1513E+02
0.1143E+03	-0.1819E+02	0.9705E+02	-0.1890E+01
-0.1399E+03	0.8153E-01	-0.4984E+00	0.9439E+00
$y = 0.05$			
-0.1765E+02	0.5144E+02	-0.5832E+02	0.7095E+02
-0.2557E+03	-0.7899E+02	0.4769E+03	0.2965E+02
-0.2393E+03	0.4745E+00	-0.1174E+01	0.6081E+01

Born term cross-section, within the allowed phase-space region. Each event is assigned a weight, given by $1 + (\alpha_S/\pi)F(X,Y;y)$. In a run, the total three-jet fraction can be obtained by averaging these weights, multiplied by the relevant Born term cross-sections and phase-space factors. Four- and two-jets are generated with weight unity. With also the four-jet fraction kept track of, the correct relative two-jet fraction may be constructed at the end of the run, according to $r_{2jet} = 1 - r_{3jet} - r_{4jet}$. Since r_{2jet} differs from the actually generated two-jet fraction, r_{2jet}^{gen}, a common weight factor r_{2jet}/r_{2jet}^{gen} is applied to all two-jets.

In practice, for each kinematical variable to be studied over a range of α_S values, it is convenient to use four separate histograms during event generation:

1. 2-jet events, with weight unity,
2. 3-jet events, with weight unity,
3. 3-jet events, with weight $F(X,Y;y)$, and
4. 4-jet events, with weight unity.

Once these histograms are available, with integrated rates (as functions of α_S) for the latter three, it is straightforward to mix the four histogram contents in the desired proportions. In particular, a global change in α_S can be trivially implemented, simply by changing the mix. A change of α_S event-by-event, as resulting from initial state QED radiation, is less easily taken into account but, since α_S varies slowly with CM energy, this is not a major problem. Also, y must be kept fixed within a run, and restricted to one of the values in the table.

Of course, it is also possible to use an algorithm along the JETSET lines (see section 3.2). It is then not necessary to use separate histograms but, on the other hand, studies for different α_S values have to be based on separate runs. Actually, since the original Zhu generator was not available, an implementation into JETSET was arranged. It is this code which has been used in the comparisons presented below.

3.2 FKSS/GKS-based Generators

The only surviving generator from the FKSS/GKS school [10,3] is the one found in JETSET. It is based on the GKS calculation, where some of the original mistakes in FKSS had been corrected. The GKS formulae have the advantage of giving the second order corrections in closed analytic form, as not-too-long functions of x_1, x_2, and the y cut-off. However, it is today recognized, also by the authors, that important terms were still missing, and that the matrix elements should therefore not be taken too seriously. There are two reasons for including the JETSET generator in this report anyway, first that it has been much used in the past and is still available, and second that the methodology chosen in JETSET has been the starting point for several other generators, to be described in the following.

The program contains parametrizations, separately, of the total first order three-jet rate, the total second order three-jet rate, and the total four-jet rate, all as functions of y (and, in a trivial manner, α_S). These parametrizations have been obtained as follows.

- The three-jet matrix element is almost analytically integrable; some small finite pieces were obtained by a truncated series expansion of the relevant integrand.
- The second order three-jet matrix elements were integrated for 40 different y cut values, evenly distributed in $\ln y$ between a smallest value $y = 0.001$ and the kinematical limit $y = 1/3$. For each y value, 250 000 phase-space points were generated, evenly in $d\ln(1-x_i) = dx_i/(1-x_i)$, $i = 1,2$, and the second order three-jet rate in the point evaluated. The sum of weights in each of the 40 y points, properly normalized, were then fitted to a polynomial in $\ln(y^{-1}-2)$.
- The four-jet rate was integrated numerically, separately for $q\bar{q}gg$ and $q\bar{q}q'\bar{q}'$ events, by generating large samples (exact size lost to history) of four-jet phase-space points within the boundary $y = 0.001$. Each point was classified according to the actual minimum y between any two partons. The same events could then be used to update the summed weights for 40 different counters, corresponding to y values evenly distributed in $\ln y$ between $y = 0.001$ and the kinematical limit $y = 1/6$. In fact, since the weight sums for large y values only received contributions from few phase-space points, extra (smaller) subsamples of events were generated with larger y cuts. The summed weights, properly normalized, were then parametrized in terms of polynomials in $\ln(y^{-1}-5)$. Since it turned out to be difficult to obtain one single good fit over the whole range of y values, different parametrizations are used above and below $y = 0.018$.

In the generation stage, each event is treated on its own, which means that the α_S and y values may be allowed to vary from event to event. The main steps are the following.

1. The y value to be used in the current event is determined. If possible, this is the value given by the user, but additional constraints exist from the validity of the parametrizations ($y > 0.001$), an extra user requirement of a minimum absolute invariant mass between jets (which translates into varying y cuts due to the effects of initial state QED radiation), and the demand that the first plus second order three-jet cross-section never be negative in any phase-space point (while the second order three-jet cross-section by itself may well be negative). The latter constraint on y, as a function of α_S, was found as a by-product of the numerical integration described above.
2. For the y and α_S values given, the relative two/three/four-jet composition is determined. This is achieved by using the parametrized functions of y for three- and four-jet rates, multiplied by the relevant number of factors of α_S, and normalized to the analytically known second order total event rate.
3. If the combination of y and α_S values is such that the total three- plus four-jet fraction is larger than unity, i.e. the remainder two-jet fraction negative, the y cut value is raised (for that event), and the process is started over at point 2.
4. The choice is made between generating a two-, three- or four-jet event.
5. For the generation of four-jets, it is first necessary to make a choice between $q\bar{q}gg$ and $q\bar{q}q'\bar{q}'$ events, according to the relative (parametrized) total cross-sections. A phase-space point is then selected, and the differential cross-section in this point is evaluated and compared with a parametrized maximum weight. If the phase-space point is rejected, a new one is selected, until an acceptable four-jet event is found.
6. For three-jets, a phase-space point is first chosen according to the first order cross-section. For this point, the weight
$$W(x_1, x_2; y) = 1 + \frac{\alpha_S}{\pi} R(x_1, x_2; y), \tag{37}$$
is evaluated, where $R(x_1, x_2; y)$ is the analytically given [3] ratio of second to first order three-jet cross-sections. This weight is compared with a maximum weight
$$W_{max}(y) = 1 + \frac{\alpha_S}{\pi} R_{max}(y), \tag{38}$$
where also $R_{max}(y)$ was obtained as a by-product of the numerical integration above. If the phase-space point is rejected, a new point is generated, etc.
7. Massive matrix elements are not available in JETSET. However, if a three- or four-jet event determined above falls outside the phase-space region allowed for massive quarks, the event is rejected and reassigned to be a two-jet event. This procedure is known not to give the full mass suppression expected, but at LEP mass effects will be minor anyway.
8. Finally, if the event is classified as a two-jet event, either because it was initially so assigned, or because it failed the massive phase-space cuts for three- and four-jets, the generation of two-jets is trivial.

The parametrizations were originally done for PETRA applications. A legacy of this is the restriction to 4 flavours for the secondary $q'\bar{q}'$ pair generated in $q\bar{q}q'\bar{q}'$ events, i.e. $b\bar{b}b\bar{b}$ events are not generated. Further, the overall angular orientation of events (with respect to the beam axis) is based on the first order electroweak formulae, both for three-jets and four-jets. For four-jets this means that they are oriented in analogy with effective three-jets obtained by recombining the gg or $q'\bar{q}'$ parton pair into one single g.

3.3 ERT-based Generators (II)

The event generator written by F. Csikor has been used e.g. in the CELLO collaboration. This generator uses an early version of the Zhu parametrization, in terms of 18 parameters rather than 12, but with a comparable or even slightly worse χ^2 for the fit. This generator has been fitted into the JETSET framework, which means that the program can be run almost in the ordinary JETSET way. Only five y values are supported: $y = 0.011, 0.02, 0.03, 0.04, 0.05$, and the program is said to be valid only for 35 and 44 GeV. In practice, it seems these constraints can be relaxed.

The selection to generate a two-, three- or four-jet event is done in the standard JETSET way, but with the second order three-jet contribution set to zero. This automatically gives the right fraction of four-jets, so these are assigned weight unity. The rate of two-jet events is too high, however, at the expense of three-jets. Two-jets are therefore given a weight

$$W_2 = \frac{1 - r_3^{(1)} - r_3^{(2)} - r_4}{1 - r_3^{(1)} - r_4}, \tag{39}$$

where $r_3^{(2)}$ is the parametrized second order contribution to the three-jet fraction. Three-jets, finally, are generated according to the first order cross-section, and then assigned a weight $1 + (\alpha_S/\pi)F'(X, Y; y)$, where F' is an 18 parameter function similar to the F of eq. (36), i.e. supposed to be an approximation to the R function of eq. (35). Once calculated, the two-, three-, or four-jet weight is compared with a maximum weight $1 + (\alpha_S/\pi)F'_{max}(y)$, where F'_{max} is again given in tabular form. If the event is rejected, a new one is generated, starting with the selection of jet multiplicity.

The parametrization is fitted into JETSET version 5.2 (but could easily be moved to 6.3), as follows. The routines LUEEVT, LUXJET and LUX3JT are replaced by modified versions, and in addition there are two extra routines ALPCOR and ERT2ND. To switch on the option, one should put MSTE(20)=2.

The approach above could probably be cleaned up by using the parametrized second order three-jet rate to make a definitive choice of jet type (apart from mass effects) from the beginning, just like in JETSET. However, in principle it is a correct method.

Of more recent date are the ERT matrix elements given by N. Magnussen (Wuppertal and DESY, BITNET F11MAG@DHHDESY3). [11]. These are intended to be used in the context of an optimized choice of Q^2 scale. As mentioned in section 2.2, results are only independent of the choice of scale if the infinite perturbative series is taken into account. Given that only two orders are at our disposal, one might e.g. try to apply the principle of minimum sensitivity, PMS [12], which states that the optimum choice of scale Q^2 is the one for which physical quantities are least sensitive to small variations of scale. Traditionally, an unoptimized scale $Q^2 = s$ has been used, but it is now established that this fails to give an adequate description of the large rate of multi-jet events observed experimentally [13]. Kramer and Lampe [14] have recently applied the PMS to jet cross-sections. They find that, by doing so, the rate of four-jet events is significantly larger than that predicted with the normal, unoptimized scale, thereby rekindling hope that matrix elements could be suitable for many tasks. Experimental studies have confirmed that the use of second order matrix elements with an optimized α_S scale results in jet rates in nice accord with data [11,15]. Moreover, many of the global event distributions are also well described. Recently, Magnussen has managed

to obtain a reasonable description also of the p_T^{out} distribution, notoriously difficult to describe with matrix elements, by using a larger-than-usual fragmentation p_T.

For the calculation of the Magnussen matrix elements, the Monte Carlo method of Ali and Barreiro was used [16]. As in the Zhu calculation, described above, part of the second order three-jet cross-section is calculated analytically, while part is obtained numerically, by recombination of four-parton events. The results were based on a sample of 6 500 000 four-parton and 1 900 000 three-parton events. Three different recombination schemes were studied:

1. the momentum one, $\overline{p}$,
2. the massless energy one, E_0, and
3. the energy one, E.

In the E_0 recombination scheme, which has not been encountered before, the four-vector of the recombined parton is constructed as the sum of the two nearby parton four-momenta. The scaled invariant mass-squares in the three-jet configuration are calculated from energies and angles, according to $y_{ij} = 2E_iE_j(1 - \cos\theta_{ij})/E_{CM}^2$, i.e. a formula that would have been correct had the recombined parton been massless. As it is, $y_{12} + y_{13} + y_{23} \neq 1$. The choice here is to have the energy fractions of the quark and antiquark given by $x_i = 1 - y_{jk}$, while the gluon takes what is left (i.e. not $1 - y_{12}$ but $2 - x_1 - x_2$).

Also the problems with unphysical recombinations of partons were considered, like the q and $\overline{q}$ of a $q\overline{q}gg$ event. The approach adopted was to include only flavour-allowed recombinations in the weight calculation for a given x_1, x_2 kinematical setup, but to include all recombinations in the calculation of the total three-jet rate.

In the end, the Magnussen algorithm is embedded in the JETSET (version 6.3) framework, by a change of three pieces of software.

- The parametrized second order three-jet cross-section is replaced by a new (similar) parametrization for each of the three recombination schemes separately.
- The ratio of second to first order weights, $R(x_1, x_2; y)$ in eq. (37), is replaced by one two-dimensional table for each y value (and recombination scheme) studied. Bin size in the (x_1, x_2) plane is 0.1×0.1. Values for the maximum weight $R_{max}(y)$ are also given. Interpolation is done for intermediate, not tabulated y values.
- The α_S function is replaced by the one relevant for an optimized Q^2 scale.

To distinguish these routines from the standard JETSET ones, the driver routine LUEEVT is instead called LUEEV1. It can be found in the file ERTLND1 FORTRAN on the JTMC 191 disk on CERNVM. Three sets of matrix elements reside in the file ERTMES1 FORTRAN. They can be selected with the JETSET switch MSTE(1).

- MSTE(1) = 4 : $\overline{p}$ recombination at scale $Q^2 = s$, $\alpha_S(0.005s)$, $\Lambda = 120$ MeV.
- MSTE(1) = 5 : E recombination at scale $Q^2 = s$, $\alpha_S(0.005s)$, $\Lambda = 50$ MeV.
- MSTE(1) = 6 : E_0 recombination at scale $Q^2 = 0.005s$, $\alpha_S(0.005s)$, $\Lambda = 90$ MeV.

The last option gives the same total three-jet cross-section as that calculated by Kramer and Lampe, but no comparison of the differential cross-section has been made.

For completeness, the main additional JETSET parameters set by Magnussen in his study at 44 GeV were the following:
MST(4) = 3 : freedom to set c and b fragmentation functions separately;

PAR(12) = 0.375 : fragmentation p_T σ parameter (0.425 in updated analysis using MSTE(1)=6);
PAR(31) = 1.0 : a parameter in Lund light quark fragmentation function;
PAR(32) = 0.6 : b parameter in Lund light quark fragmentation function;
PAR(44) = -0.050 : ϵ parameter for c fragmentation function;
PAR(45) = -0.018 : ϵ parameter for b fragmentation function; and
PARE(8) = 0.015 : y cut-off of matrix elements.

When running this program, small jumps in the parton thrust distribution, in particular at $T = 0.9$, were noted. These come from the use of a fairly coarse matrix of second order correction factors, $\Delta x_1 \times \Delta x_2 = 0.1 \times 0.1$, with no smoothing at boundaries.

3.4 GS-based Generators

The Gottschalk-Shatz calculation [17] was an attempt to perform a completely analytical calculation of the second order three-jet rate, by integrating the ERT four-jet matrix elements in the regions below a given physical y cut-off scale. A 'partial fractioning' scheme was used to subdivide the matrix elements by their pole structure, and perform analytic cancellation one pole at a time. Parton recombination was done according to the E scheme, i.e. the energies of the two recombined partons were added. Unfortunately, the calculations were never completed, in the sense that not all soft/collinear terms could be integrated analytically. Therefore, in the matrix element package put together by Gottschalk, some of the generated four-jet events have two partons with an invariant mass smaller than that allowed by the y cut-off. Since the singularities have been cancelled completely, the cross-section in the soft four-jet region is still reasonably small, however. If the fragmentation algorithm can handle them, these events can therefore be generated as they are. Alternatively, they can be recombined according to some scheme; whether $\overline{p}$ or E is used in that case is easily seen to make essentially no difference at all.

In the Gottschalk package, five event classes are distinguished, and treated separately:

1. two-jets,
2. three-jets,
3. four-jets with a $q^* \to qg$ pole (i.e. Abelian type $q\overline{q}gg$),
4. four-jets with a $g^* \to gg$ pole (i.e. non-Abelian type $q\overline{q}gg$), and
5. four-jets with a $g^* \to q'\overline{q}'$ pole (i.e. $q\overline{q}q'\overline{q}'$).

Due to the somewhat special nature of the cancellation procedure, the y cut-off can be specified separately for four-jet and three-jet events; in the program the former is called y and the latter ϵ. The third input parameter is α_S. With these three given, an initialization routine may be called, which will determine the fractions of event classes 2 - 5 above by Monte Carlo integration, with the two-jet rate given as the remainder. Simultaneously, maximum weights are also determined. Values obtained at initialization are listed, and may be directly input in a subsequent run with the same parameter values. A third initialization option exists, where the authors parametrize fractions and maximum weights for a range of standard parameter values.

In the event generation, a choice is made from the start between one of the five event classes. Phase-space points are then generated according to the pole structure, the full weight is evaluated in the point chosen, and this weight is compared with the

(previously determined) maximum weight. If the event is rejected, a new one of the same class is generated. Quark mass effects are never considered at all.

Two implementations have been studied. One is by S. Bethke (BITNET SIGGI@ CERNVM). Here some further pretabulation has been performed, so that the average user need never worry about the initialization stage. Further, the events generated may be interfaced into JETSET. For this purpose, a modified version of LUEEVT is available, so that generation can be run in standard fashion. Four-jet events with two nearby partons (i.e. which fail the y cut but were still kept by GS) are handed on directly to the fragmentation routines. This program may be found in the file GSME FORTRAN on the TORSJO 192 disk on CERNVM. An additional routine for optional recombination of two nearby partons in four-parton events has been added.

Another program based on the GS matrix elements has been written by R.-y. Zhu and H. Stone (obtainable via BITNET, STONE@VXCRNA.CERN.CH) [18,19]. In this program, the GS matrix elements may be evaluated (by Monte Carlo methods) for a cut-off scale y_0, where $y_0 \leq y$. Any four-parton events failing the physical cut-off scale y are recombined into three-jet ones, using the $\overline{p}$ recombination scheme. Typically, a $y_0 = 10^{-3}$ is used.

3.5 KL-based Generators

A complete recalculation of the second order three-jet differential cross-section has been performed by Kramer and Lampe [20,14]. Like in the Gottschalk-Shatz calculation above, a partial fractioning scheme is used, i.e. the four-jet matrix element is rewritten as a sum of terms, in such a way that each term is singular in at most one scaled invariant mass y_{ij}. Such a term is integrated analytically in the region $y_{ij} < y$, where it is singular. In other regions, like $y_{kl} < y_{ij}$, $kl \neq ij$, the integration is performed numerically, to give finite $\mathcal{O}(y)$ terms.

As with other schemes, there is an ambiguity in the choice of recombination scheme to project four-jet kinematics onto three-jet kinematics. Two different schemes have been compared, called KL and KL'. The authors tend to prefer the latter one [14], KL'. This scheme is essentially the same as the E scheme, in that the energy of a recombined parton is obtained as $E_{ij} = E_i + E_j$, with three-momenta suitably modified. In the study of a term singular in y_{ij}, in the region $y_{ij} < y$, it is always the partons ij that are recombined, even if another parton pair happens to have a smaller invariant mass, e.g. if $y_{ik} < y_{ij}$. For the numerical integration in other regions, where the term is not singular anyway, the energy recombination is performed for the pair with small invariant mass in that region.

In the other scheme, the KL one, if the singular term y_{ij} corresponds to a qg pair, the energy of the $\overline{q}$ is unchanged in the recombination procedure, while the energy of the unrecombined gluon is increased by $y_{ij}E_{CM}/2$, and that of the pair decreased by the corresponding amount. If the recombined pair is gg of $q'\overline{q}'$, the energy of the antiquark is unchanged, while that of the quark is increased as above. The procedure is thus not symmetric, but that could probably be fixed trivially by symmetrizing the final result. In general, second order corrections tend to be larger in the KL scheme than in the KL' one.

The Kramer-Lampe calculations were not immediately implemented in a fully-fledged Monte Carlo event generator, due to the complexity and length of the formulae involved.

Results have been given for jet rates as a function of the y cut-off, in particular when the Q^2 scale is optimized [14]. Work, by Kramer and Magnussen, has been under way to write a user-friendly event generator based on the KL matrix elements and the use of an optimized Q^2 scale. Just as this manuscript was finalized, we learned that the program is now ready. Time has not permitted us to look at the program, but we expect it will provide a very valuable addition to the other generators already described.

3.6 Comparison of Generators

Some comparisons of event generators exist in the literature [6,11,18,19]. In particular, good agreement has been demonstrated between the Zhu and GS programs (more about this later). However, most comparisons have only involved two programs at a time. Therefore, several of the event generators described above have been test run, so as to compare results. In order not to introduce extra complications, there is no initial state QED radiation, quarks are always massless, and α_S is fixed at 0.12. Since absolute cross-sections were never considered, but only the relative composition of events, results do not depend on the actual energy scale used; for convenience $E_{CM} = m_Z = 92$ GeV was assumed. Two different cut-off values were used, $y = 0.01$ and $y = 0.05$. Results are given in Table 2, for the first order three-jet fraction $r_3^{(1)}$, the second order additional contribution $r_3^{(2)}$, such that the total three-jet fraction is $r_3 = r_3^{(1)} + r_3^{(2)}$, and the four-jet fraction r_4.

Note that the agreement in $r_3^{(1)}$ and r_4 is not as impressive as it may look: all generators based on JETSET here make use of the same parametrizations. Also, where $r_3^{(1)}$ and $r_3^{(2)}$ are not given separately, the JETSET $r_3^{(1)}$ value is assumed and used to derive $r_3^{(2)}$. The numbers quoted are the ones contained in these parametrizations, where given, or else the results obtained in runs with 500 000 events per generator (in some cases somewhat less).

The ERT(Zhu) figures are the ones obtained with the parametrization in section 3.1, and the ones obtained with the original Zhu program, respectively. Only the parametrization has been used in the following.

One should note that the Magnussen program was really intended for use with an optimized Q^2 scale; we have here disregarded that possibility, in order to allow comparisons with the other programs. The comparison may therefore be slightly unfair. For the E recombination scheme, the roof is actually hit, where naively $r_3^{(1)}+r_3^{(2)}+r_4 \geq 1$; this is reflected in a reduced four-jet rate as the program tries to cope. Magnussen has supplied us with numbers, where calculations were made without reference to the crude binning of R in the public program, and where some missing non-singular terms have been included. For $y = 0.01$, the $r_3^{(2)}$ values are then typically reduced by 0.01.

The GS(Stone) alternative is based on using a $y_0 = 10^{-3}$ in the GS matrix elements, and then recombining partons up to the physical y scale. As it turns out, results are rather sensitive to the choice of y_0 value: if y_0 is decreased, so is also $r_3^{(2)}$ (while r_4 stays constant, as it should). For $y = 0.01$, the variation is from $r_3^{(2)} = +0.100$ for $y_0 = y = 0.01$ to -0.017 for $y_0 = 10^{-5}$. This variation could be a reflection of the mismatch between the E recombination scheme used in the GS matrix elements and the $\overline{p}$ one of the Stone recombination step; indeed the difference in $r_3^{(2)}$ between the two extreme choices of y_0 is comparable to the split seen between the Magnussen E and $\overline{p}$ recombination schemes (although the absolute level is shifted). The previously

Table 2: Relative jet composition for second order matrix element event generators, with masses neglected and $\alpha_S = 0.12$.

y	generator	$r_3^{(1)}$	$r_3^{(2)}$	r_4
0.01	ERT (Zhu, parametr.)	0.684	0.081	0.116
	ERT (Zhu, full prog.)	0.684	0.072	0.118
	GKS (JETSET)	0.684	-0.032	0.116
	ERT (Csikor)	0.681	0.026	0.109
	ERT (Magnussen, $\overline{p}$)	0.684	0.073	0.116
	ERT (Magnussen, E_0)	0.684	0.096	0.116
	ERT (Magnussen, E)	0.684	0.200	0.105
	GS (Bethke)	0.684	0.069	0.125
	GS (Stone)	0.684	0.100	0.131
	GS (Zhu)	0.684	0.068	0.131
	KL′ (Magnussen)	0.684	0.097	0.116
0.05	ERT (Zhu, parametr.)	0.212	0.067	0.0047
	ERT (Zhu, full prog.)	0.212	0.061	0.0047
	GKS (JETSET)	0.212	0.049	0.0047
	ERT (Csikor)	0.211	0.067	0.0046
	ERT (Magnussen, $\overline{p}$)	0.212	0.073	0.0047
	ERT (Magnussen, E_0)	0.212	0.081	0.0047
	ERT (Magnussen, E)	0.212	0.111	0.0047
	GS (Bethke)	0.212	0.088	0.0050
	GS (Stone)	0.212	0.075	0.0054
	GS (Zhu)	0.212	0.058	0.0054
	KL′ (Magnussen)	0.212	0.084	0.0047

mentioned agreement between the ERT(Zhu) and the GS results is therefore probably only a coincidence, and could be destroyed by another choice of y_0.

Inclusive jet rates obviously do not give the full story. In particular, the r_3 values are dominated by the behaviour close to the cutoff at thrust $T = 0.99$ (or 0.95), while the ordinary event shape studies effectively involve weights that favour the low-thrust events. Since everybody should agree on the shape of four-jets, we have looked at the three-jet thrust distribution as the next simplest differential quantity. The result for $y = 0.01$ is shown in Fig. 11, with the first plus second order three-jet thrust distribution normalized to the (well-known) first order one. Original results were for runs with 500 000 events and for bins of size 0.01, and were then smoothed by hand to even out statistical fluctuations.

The results for the GS (Zhu/Stone) programs are not included in the figure; separately it has been shown that these programs agree very well with the ERT(Zhu) results [18,19] but, as noted above, this might be due to a fortuitous choice of y_0 value. Neither have we had access to the KL′(Magnussen) results for the thrust distribution.

While the GKS curve is known to be wrong, there is enough of a spread between the other ones. This spread is not seen so much in the inclusive jet rate, i.e. $r_3^{(2)}$, since that rate is dominated by the region close to thrust $T = 1$, where the spread between the programs tends to be smaller. Experimental α_S determinations are typically based

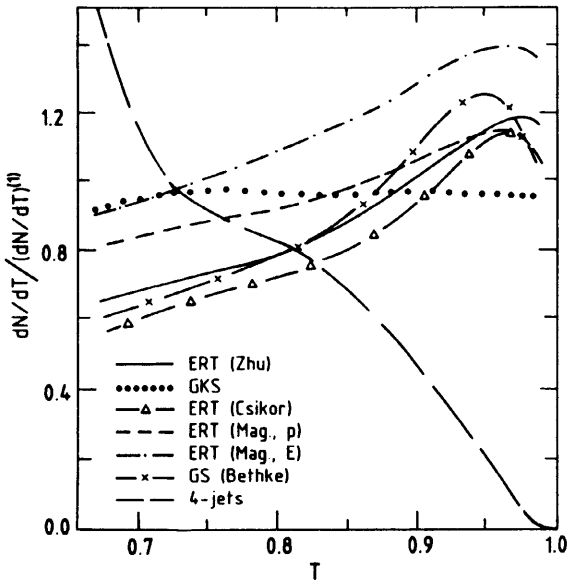

Figure 11: Full second order three-jet thrust distribution, normalized to the first order one, for $\alpha_S = 0.12$ and $y = 0.01$. For additional details, see text.

on the shape of the thrust distribution, and would thus also be sensitive to the region $T < 0.85$. Here, however, four-jets give a non-negligible contribution — actually the ratio becomes infinite at $T = 2/3$, since three-jet phase-space vanishes at that point, while four-jets may have $T < 2/3$. The net effect is thus that the uncertainty in the value of the thrust distribution is nowhere larger than $\pm 10\%$, based on the generators we have studied. Currently, this is therefore a realistic error to assign to any α_S determination, just from the matrix element point of view. Additional uncertainties come from higher orders, from fragmentation, from experimental errors, and so on.

3.7 Four-jet Matrix Elements

Many programs have four-jet matrix elements available. The list includes two of the programs of the next section, the Hagiwara-Zeppenfeld and the Berends-Giele-Kuijf ones, as well as e.g. JETSET. These three all assume quarks to be massless, however. Furthermore, as we have noted previously, the $q\bar{q}q'\bar{q}'$ final state receives contributions from graphs with no analogue in pure γ exchange [5].

A program for $Z^0 \to q\bar{q}q'\bar{q}'$ decays, with full quark mass dependence and angular orientation (w.r.t. the beam axis) included, has been written by Glover, Kleiss and van der Bij [21]. It is assumed that the two quarks q and q' are not the same, since then additional graphs should be included. Contributions from γ exchange are not included, so the program is strictly intended for decays at the Z^0 peak.

After an initialization call, where the q and q' flavours have to be specified, one event can be generated at a time by a call to the main routine. This routine will return the four-momenta of q, $\bar{q}$, q' and $\bar{q}'$, and an associated event weight. The program can be

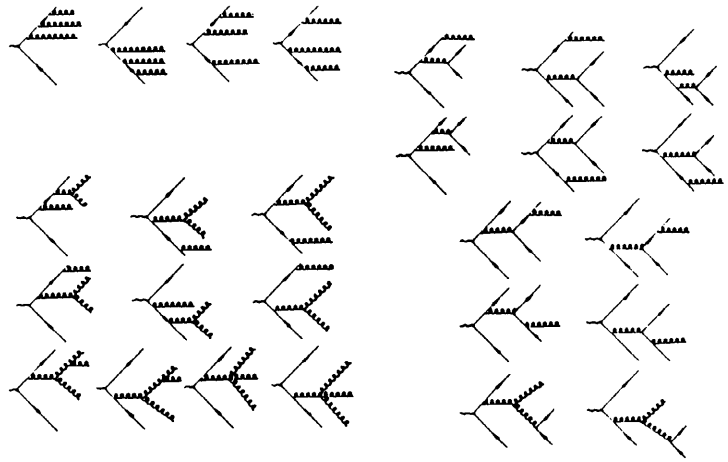

Figure 12: Feynman diagrams for the production of five partons at $\mathcal{O}(\alpha_S^3)$. All permutations of the gluon and the quark-antiquark lines have to be included.

found in the file ZFFFF FORTRAN on the CERNVM JOCHUM 191 disk.

3.8 Five-jet Matrix Elements

Recently, three different groups [22,23,24] have independently calculated all tree-level matrix elements contributing to the processes $e^+e^- \to q\bar{q}ggg$, $q\bar{q}q'\bar{q}'g$ at $\mathcal{O}(\alpha_S^3)$. They all assume massless partons, although they keep open the possibility to evaluate the corresponding cross-sections with massive quarks. The whole set of Feynman diagrams giving rise to these final states is shown in Fig. 12.

The results of these calculations are available in programs, which can be used to generate one parton configuration at a time. The programs are selfcontained, and any interfacing to a fragmentation program for partons has to be done by the user himself. Here below we shall briefly discuss the three programs and make some comparisons. In fact, the three groups have already checked their results against each other, by picking a number of specific parton configurations, for which the cross-sections were evaluated numerically. The same answers were then obtained, up to machine precision.

In the method of Berends, Giele and Kuijf, use is made of recursive relations for calculating the matrix elements

$$\mathcal{M}(q_1, q_2; k_1, \ldots, k_n) = V^\mu \hat{S}_\mu(q_1, q_2; k_1, \ldots, k_n) \tag{40}$$

for $e^+e^- \to q\bar{q} + ng$, and

$$\mathcal{M}(q_1, q_2, q_3, q_4; k_1, \ldots, k_n) = V^\mu \hat{T}_\mu(q_1, q_2, q_3, q_4; k_1, \ldots, k_n) \tag{41}$$

for $e^+e^- \to q\bar{q}q\bar{q} + ng$. In these expressions, V^μ is the polarization vector for the vector boson, whereas $\hat{S}_\mu$ and $\hat{T}_\mu$ are currents, containing quarks and gluons. These currents depend on the momenta $k_1, .., k_n$ of the outgoing gluons, the momenta q_1, q_3 (q_2, q_4) of the outgoing quarks (antiquarks), the helicities and the colours of the partons.

The essential point of their method is that the recursion relations calculate a current, i.e. an amplitude, in which one parton is off-shell. In order to get the physical amplitude it has to be contracted with a polarization vector (for an off-shell gluon) or a spinor (for an off-shell quark/antiquark) and finally, momentum conservation has to be imposed. In this way the colour structure is systematically developed by the recursion relation, and the Feynman diagrams are automatically generated: for an n-gluon process, building blocks can be used which originated from processes with $n-1$ gluons, and so on.

The program of Berends, Giele and Kuijf is based on a main routine, which initializes all Standard Model parameters and physical cuts for the generation. It calls the VEGAS package, which integrates the appropriate cross-section. The program considers all dynamical and kinematical factors and finally produces some output with the result of each step, including the evaluation for the final state cross-section. In order to make distributions of any variable, each generated event kinematics has to be weighted with the corresponding matrix elements.

The Hagiwara and Zeppenfeld matrix elements are based on a notation which makes it easy to obtain the amplitude of a particular process by crossing from a generic amplitude. They are valid either for real or virtual vector bosons. For 4-parton (or 5-parton) amplitudes one employs an orthogonal basis of the colour factors, by using the representation matrices of the permutation groups S_2 and S_3. This way the authors claim to minimize the number of Feynman diagrams to be evaluated, and to make the sum over colours trivial and the treatment of identical particles easy. For fixed polarizations of the external particles, the matrix elements can be written in a general form:

$$\mathcal{M} = \sum_m M^{(m)} O^{(m)} \Rightarrow \sum_{colours} |\mathcal{M}|^2 = \sum_m c_m |M^{(m)}|^2 \qquad (42)$$

where $O^{(m)}$ are the orthogonal colour tensors and c_m are some tabulated factors (see tables 1, 2, 3 and 4 from their report [22]).

The structure of the program is analogous to the preceding one; also here a weighting procedure has to be employed in order to get correct distributions. However, in this case the user himself has to provide a main program containing a phase-space generator; when fed with a parton configuration the routines will return the colour summed, polarization summed/averaged matrix element squared. Moreover, the program has two alternative running modes, either with a random (R) or an explicit (S) polarization summation.

In the study group, we have made a simple comparison of the two programs. The differential cross-section $d\sigma/dy_{min}$ was calculated, where y_{min} is the scaled minimum invariant mass-squared between any two partons in the event. The parameters for the generation were $\alpha = 1/137$, $\alpha_S = 0.133$, $\sin^2\theta_W = 0.23$, $m_Z = 92$ GeV, $\sqrt{s} = m_Z$, and the cut in the minimum scaled invariant mass was $y = 0.01$. The result is shown in Fig. 13. We conclude that an excellent agreement is observed for the two programs.

Finally the Falck, Graudent and Kramer group has calculated the trace for the five-jet cross-sections by using REDUCE 3.2, with traditional methods, in the Feynman gauge. Since all terms in the matrix elements are explicitly evaluated (rather than appearing implicit in an recursive procedure, e.g.), the actual code is rather longer than that of the other two programs.

In summary, jets will become better resolvable at LEP than they were at lower energies, and thus we expect these programs to become very useful for testing QCD (triple-gluon-vertex, topological studies, etc.), or for making reliable estimates of the

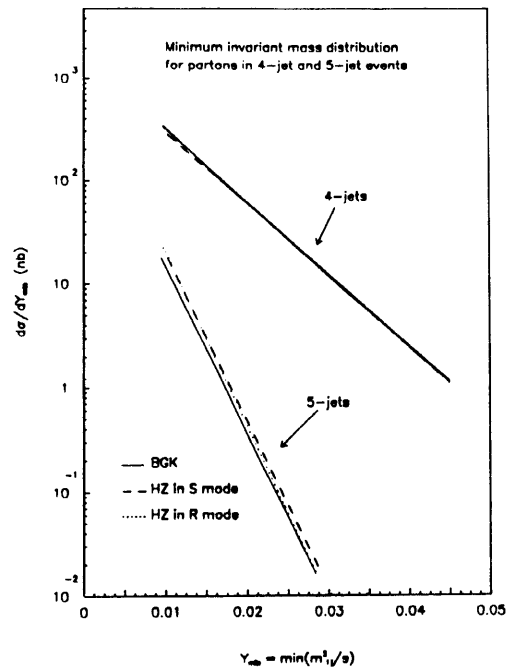

Figure 13: Good agreement is found for the prediction of the 4-jet and 5-jet cross-sections in the programs from Berends, Giele and Kuijf (BGK) [24] and Hagiwara and Zeppenfeld (HZ) [22], the latter in its two run modes, S and R, as described in the text. The small differences which can be seen in the figure are consistent with being statistical fluctuations.

QCD backgrounds to any new phenomena with a multi-jet final state signature.

All programs can be obtained by request to the authors, e.g. BITNET LORENTZ@HLERUL5 for Giele, DIETER@WISCPHEN for Zeppenfeld, and I02KRA@DHHDESY3 for Kramer. The first two programs can also be found in the files XSE5BGK FORTRAN and XSEC5HZ FORTRAN on the FUSTER 191 disk on CERNVM.

3.9 Multi-jet Approximation Techniques

In the previous sections, we have seen that several matrix elements now are available for multiparton events. As the number of partons is increased, the resulting expressions can become quite complicated, however. Thus, even when they are known, they may be rather time-consuming to evaluate. An alternative approach, based on approximate methods that gives much simpler (but still surprisingly accurate) formulae, has been pursued by C. J. Maxwell (Department of Physics, Durham University, South Road, Durham DH1 3LE, United Kingdom, BITNET CJM@HEP.DURHAM.AC.UK). The story goes as follows.

The two processes $e^+e^- \to q\bar{q}$+gluons, and $e^+e^- \to q\bar{q}q'\bar{q}'$+gluon(s) have now been calculated for up to five final jets [22,23,24]. The former process is dominant, providing about 95% of the total multi-jet rate for four jets, and around 85% for five jets. One can construct a simple approximation to the matrix element squared of this dominant sub-process, which would enable six- or more-jet production to be estimated at LEP. The approximation agrees at the 10% level with the known exact results for four- and

five-jets, and is exact for three-jet production.

We shall summarize very briefly how the approximation is constructed. Full details can be found in ref. [25].

The approximation is based on the remarkable fact that, for $e^+(p_1) + e^-(p_2) \to q(q_1)+\bar{q}(q_2)+g(k_1)+\ldots+g(k_n)$, one can write down the tree-level helicity sub-amplitude squared for all gluons having the same helicity, to leading order in the number of colours, N_C, for an arbitrary number of gluons, n [26].

One has

$$\begin{aligned} |\mathcal{M}_n^S|^2 &= e^4 g^{2n} N_C^{n+1} 2^{3-n} s\{C_{RR}[(p_1q_1)^2 + (p_2q_2)^2] + C_{RL}[(p_1q_2)^2 + (p_2q_1)^2]\} \\ &\quad \times \sum_P \frac{1}{(q_1k_1)(k_1k_2)\ldots(k_nq_2)}. \end{aligned} \tag{43}$$

The result includes both photon and Z^0 exchange. (p_iq_i) is used to denote the dot product of four-momenta, $p_i \cdot q_i$. 'P' denotes a sum over all n! permutations of the gluon momenta $k_1, k_2, \ldots, k_n$. C_{RR} and C_{RL} are various combinations of electroweak charges.

Denoting the full amplitude squared for this process as $|\mathcal{M}|^2$, one finds that, up to a trivial factor due to the leading N_C approximation, the three-jet result is exactly given by the sub-amplitude (43),

$$|\mathcal{M}_1|^2 = \frac{8}{9}|\mathcal{M}_1^S|^2. \tag{44}$$

To construct an approximation for $n > 1$, one uses the fact that the ratio $|\mathcal{M}_n|^2/\mathcal{M}_n^S|^2$ contains no kinematical poles and is hence rather insensitive to the exact momentum configuration for which it is evaluated. Replacing the exact momentum configuration by one in which the smallest invariant mass pair of jets, i, j, are replaced by a collinear pair in the direction of $\bar{p}_i + \bar{p}_j$, with energy fractions $z = E_i/(E_i + E_j)$ and $(1-z)$, one finds that

$$\frac{|\mathcal{M}_n|^2}{|\mathcal{M}_n^S|^2} \simeq F(R,z)\frac{|\mathcal{M}_{n-1}|^2}{|\mathcal{M}_{n-1}^S|^2}. \tag{45}$$

Here R is a simple function of the momenta after their 'reduction', and

$$F(R,z) = \begin{cases} z^4 + (1-z)^4 + 1 & i,j = g,g \\ \frac{8}{9}\frac{(1+R)(1+z^2)}{R+z^2} & i,j = q,g \end{cases} \tag{46}$$

Applying eq. (45) recursively one finally obtains the approximation

$$|\mathcal{M}_n|^2 \simeq \frac{8}{9}\prod_{i=1}^{n-1} F(R_i, z_i)|\mathcal{M}_n^S|^2 \tag{47}$$

where the factor of $|\mathcal{M}_n^S|^2$ is evaluated for the exact $(n+4)$-particle (e^+, e^-, q, $\bar{q}$, n gluons) configuration.

One can compare this simple approximation with the exact four- and five-jet results of [22,24]. One finds, applying a simple cut y, on the minimum scaled invariant mass squared of any pair of jets, $M_{ij}^2/s \geq y$, and working at two CM energies, $\sqrt{s} = 35$ GeV and $\sqrt{s} = 125$ GeV, below and above the Z^0 mass, that the approximate and exact integrated cross-sections agree at the 10% level, whilst the approximation is typically within 30% of the exact result over most of phase-space.

The success of the approximation is due to the fact that the special sub-amplitude squared of eq. (43) contributes a considerable fraction of the final result, just under half for $n = 3$ for instance. Thus the correction factors multiplying $|\mathcal{M}_n^S|^2$ in eq. (47), are of moderate size and smoothly-behaved.

Similar approximations can be constructed for a variety of multi-parton scattering sub-processes [27]. It is worth noting that in the e^+e^- case, the approximation saves a factor ~ 20 in CPU time per event, compared with the exact code [22,24].

It is evidently more than accurate enough for tree-level phenomenology, since far greater uncertainties will result from the factor $\alpha_S^n(Q^2)$, with its unknown Q^2 scale.

It would be of interest to study the relationship between the tree-level cross-sections for producing some number of jets, and the corresponding rates predicted from parton shower Monte Carlo programs. This comparison is complicated by the fact that the Born cross-section has no dependence on jet definition, and by the presence of Sudakov form factors in the shower picture.

The program is obtainable on request from the author. A copy also exists in the file ELPOSA FORTRAN on the TORSJO 192 disk on CERNVM.

3.10 QCDLIB

The QCDLIB package is not an event generator, but rather a program for the evaluation of α_S for a given Λ, or the other way around. As such, it is maybe a bit peripheral compared to other programs mentioned here, but several of the tasks this program can perform are frequently recurring ones in data analysis, and so the package may come in handy.

The program can be used either for first order or second order QCD; the α_S expressions used are those of eq. (11) or eq. (12), respectively. At flavour thresholds, the CWZ prescription is used [28], i.e. the number of effective flavours (n_f in the α_S expressions) is supposed to change discontinuously exactly at $Q = m_q$, but with a corresponding discontinuous change in the Λ value, such that α_S itself is continuous at flavour thresholds. For first order QCD, this matching is trivial to achieve, but QCDLIB also contains a root finder for giving the various Λ values needed in the second order α_S.

The program can be set up, either by giving the Λ relevant for a certain number of effective flavours, or by giving the α_S at a certain Q scale. In addition, a choice has to be made between first and second order formulae. The program then calculates the Λ values relevant for each Q range (between mass thresholds), and can afterwards be interrogated for α_S and Λ at any desired Q scale. A few other utilities are also available, some of them relevant for e^+e^- applications, others only of interest for structure function studies.

The program is written by J. Collins, P. Johnson, S. Qian and W.-K. Tung. The code can be obtained from the authors, all at Illinois Institute of Technology (e.g. BITNET ZAQ@FNAL for Tung). A copy may also be found in the file QCDLIB FORTRAN on the TORSJO 192 disk on CERNVM.

References

[1] P. Hoyer, P. Osland, H.G. Sander, T.F. Walsh, P.M. Zerwas, Nucl. Phys. **B161**

(1979) 349

[2] A. Ali, J.G. Körner, G. Kramer, J. Willrodt, Nucl. Phys. **B168** (1980) 409; A. Ali, E. Pietarinen, G. Kramer, J. Willrodt, Phys. Lett. **B93** (1980) 155

[3] F. Gutbrod, G. Kramer, G. Schierholz, Z. Physik **C21** (1984) 235

[4] J. K. Körner, G. A. Schuler, Z. Physik **C26** (1985) 559, DESY 87-124 (1987)

[5] B.A. Kniehl, J.H. Kühn, Phys. Lett. **B224** (1989) 229

[6] R.-y. Zhu, Ph. D. Thesis (M.I.T.), MIT-LNS Report RX-1033 (1983); Caltech Report CALT-68-1306; in Proceedings of the 1984 DPF conference, Santa Fe, p. 229; in Proceedings of 1985 DPF conference, Oregon, p. 552

[7] R. K. Ellis, D. A. Ross, A. E. Terrano, Phys. Rev. Lett. **45** (1980) 1226, Nucl. Phys. **B178** (1981) 421

[8] Z. Kunszt, Phys. Lett. **B99** (1981) 429, Phys. Lett. **B107** (1981) 123

[9] A. Ali, Phys. Lett. **B110** (1982) 67

[10] K. Fabricius, G. Kramer, G. Schierholz, I. Schmitt, Phys. Lett. **B97** (1980) 431, Z. Physik **C11** (1982) 315

[11] N. Magnussen, Ph.D. Thesis, University of Wuppertal WUB-DI 88-4 and DESY F22-89-01 (1989)

[12] P. M. Stevenson, Phys. Rev. **D23** (1981) 2916, Nucl. Phys. **B203** (1982) 472, Nucl. Phys. **B231** (1984) 65

[13] JADE Collaboration, W. Bartel *et. al.*, Z. Physik **C33** (1986) 23; TASSO Collaboration, W. Braunschweig et al, Phys. Lett. **B214** (1988) 286

[14] G. Kramer, B. Lampe, Z. Physik **C39** (1988) 101

[15] CELLO Collaboration, H. J. Behrend *et. al.*, Contribution to the XXIV International Conference on High Energy Physics, Munich, 1988; S. Bethke, LBL-26958 (1989)

[16] A. Ali, F. Barreiro, Phys. Lett. **B118** (1982) 155

[17] T. D. Gottschalk, M. P. Shatz, Phys. Lett. **B150** (1985) 451, Caltech Preprints CALT-68-1172, CALT-68-1173 Z. Physik **C34** (1987) 497

[18] H. Stone, CALTECH Ph.D. Thesis (1988)

[19] R.-y. Zhu, in preparation

[20] G. Kramer, B. Lampe, Fortschr. Physik **37** (1989) 161

[21] E.W.N. Glover, R. Kleiss, J.J. van der Bij, in preparation

[22] K. Hagiwara, D. Zeppenfeld, Nucl. Phys. **B313** (1989) 560

[23] N.K. Falck, D. Graudenz, G. Kramer, Phys. Lett. **B220** (1989) 299, DESY 89-046 (1989)

[24] F.A. Berends, W.T. Giele, H. Kuijf, Leiden preprint 98-0055 (1988)

[25] C.J. Maxwell, Durham University preprint DTP/89/08 (1989)

[26] M. Mangano, Nucl. Phys. **B309** (1988) 461; F.A. Berends, W.T. Giele, Nucl. Phys. **B306** (1988) 759

[27] C.J. Maxwell, Phys. Lett. **B192** (1987) 190, Nucl. Phys. **B316** (1989) 321; M. Mangano, S.J. Parke, Phys. Rev. **D39** (1989) 758

[28] J.C. Collins, F. Wilczek, A. Zee, Phys. Rev. **D18** (1978) 242

4 Event Generators

This section contains a survey of all $e^+e^- \to$ hadrons event generators that we found, and were not already presented in section 3. They span a wide range. In one end we find general purpose event generators, including primary electroweak quark production, shower evolution, fragmentation and decay. In the other, there are programs dedicated to a specific task, like shower evolution, or containing a simplified 'toy model' more intended to illuminate specific physics aspects than to offer the best possible description to all data. To help give an overview, Table 3 contains a list of what aspects of the event generation process are considered in the various programs. Note that this is *not* intended as a quality grading.

Table 3: Physics components included in the programs covered in this section. A '•' signifies that a task is handled in the program, a '+' that it is obtained by calling on another program (of the ones listed here), a '*' that it is in preparation (but not yet available), and a '-' that it is not considered at all. For some programs, the classification may not always be trivial, so this table should only be taken as a first indication. The symbols contain no grading on the quality of a program for a given task.

Program (comment)	Hard interaction	Initial γ rad.	Parton shower (or matrix el.)	Fragmentation	Decays
ARIADNE	+	+	•	+	+
CALTECH-II	•	-	•	•	•
COJETS	*	*	•	•	•
DPSJET (non-QCD)	-	-	•	•	-
EPOS (non-QCD)	•	-	•	•	•
EURODEC	-	-	*	•	•
HERWIG	•	-	•	•	•
JETSET	•	•	•	•	•
NLLJET	•	•	•	+	+
PARJET	•	-	•	•	•
TIPTOP (heavy ferm.)	•	•	-	+	+

4.1 ARIADNE

4.1.1 Basic Facts

Program name: ARIADNE, a Monte Carlo for QCD Cascades in the Colour Dipole Formulation [1,2,3]

Version: ARIADNE 3.0 (with JETSET 6.3), from June 1989
ARIADNE 3.1 (with JETSET 7.1), from June 1989

Authors: Ulf Pettersson and Leif Lönnblad
Department of Theoretical Physics
University of Lund, Sölvegatan 14A
S-223 62 LUND, Sweden
BITNET THEP @ SELDC52
with theoretical input from G. Gustafson

Program size: 2019 lines

4.1.2 Physics Introduction

The perturbative QCD cascade can be formulated in two complementary ways, either in terms of quarks and gluons (as e.g. in HERWIG and JETSET below) or in terms of colour dipoles [4,5]. When a gluon is emitted from a dipole (e.g. from a $q\bar{q}$ pair in e^+e^- annihilation), the subsequent emission of a softer gluon is given by two independent dipoles, one stretched from the quark to the gluon and the other from the gluon to the antiquark [6].

This result is generalized in the ARIADNE approach, so that the emission of one more gluon is given by three independent dipoles, etc. In this way the angular ordering and the ordering in transverse momentum are automatically taken into account. These dipoles form links in a chain, such that one gluon connects two dipoles and one dipole connects two gluons. The gluons are ordered in the dipole chain according to their colour. This is in close correspondence with the Lund string model [7], where gluons act as transverse excitations on a stringlike field. The breaking of a dipole into two dipoles thus corresponds to one more kink on the string.

A correction term with the relative magnitude $1/N_C^2$ corresponding to radiation between the quark and the antiquark is neglected here.

The confined colour field can be described by a vortex line, similar to the magnetic field in a type II super-conductor [8]. The field of such a vortex line is the same as the field of a chain of dipoles lined up along the vortex [4].

Within the above picture,the colour dipole approximation of QCD cascades seems a 'natural' formulation. Furthermore, when constructing a Monte Carlo program, the colour dipole formulation is more convenient than that in terms of coherent parton branching, since the dipoles radiate essentially independently of each other and the azimuthal dependence is automatically treated in a proper way.

Version 1 of ARIADNE was mainly intended for e^+e^- annihilation where both the quark and the antiquark are pointlike. ARIADNE-2 extends the program to leptoproduction and hadronic collisions. In these events one end of the string (the struck quark) has participated in the hard interaction and may be regarded pointlike, while the other end of the string (i.e. the remainder of the hadron) is an extended object. This leads to a suppression of gluon emission in the region of the remainder, unless the transverse

wave length of the gluon is of the same order as the transverse extension of the emitting object [9].

ARIADNE-3, in addition, includes $q\bar{q}$ emission in the dipole cascade.

When the perturbative cascade has been generated, it may be fragmented with JETSET 6.3 or 7.1 (see section 4.8). The string fragmentation is infrared stable in the sense that the presence of soft gluons does not affect the final hadronic state.

4.1.3 Detailed Physics Description

Electroweak cross-section: see JETSET

Radiative corrections: see JETSET

Relative flavour composition: see JETSET

Perturbative QCD approach:
a) Gluon radiation. There are three different types of dipole: the dipoles stretched between a quark and an antiquark (or diquark), between a quark and a gluon, and between two gluons. The calculation of the differential cross-section for gluon emission from the dipoles is presented in [1]. The result is:

$$q\bar{q} \to q\bar{q}g: \qquad \frac{1}{\sigma}\frac{d\sigma}{dx_1dx_2} = \frac{2\alpha_S}{3\pi}\frac{x_1^2+x_3^2}{(1-x_1)(1-x_3)}, \tag{48}$$

$$qg \to qgg: \qquad \frac{1}{\sigma}\frac{d\sigma}{dx_1dx_2} = \frac{3\alpha_S}{4\pi}\frac{x_1^2+x_3^3}{(1-x_1)(1-x_3)}, \tag{49}$$

$$gg \to ggg: \qquad \frac{1}{\sigma}\frac{d\sigma}{dx_1dx_2} = \frac{3\alpha_S}{4\pi}\frac{x_1^3+x_3^3}{(1-x_1)(1-x_3)}. \tag{50}$$

Here x_1 (x_3) is the usual energy fraction $2E/W$ for the endpoint quark (antiquark, gluon), where W is the invariant mass of the dipole. A gluon corresponds to a term x^3 and a quark (antiquark) corresponds to a term x^2 in the numerator of eqs. (48) - (50).

The ordinary Altarelli-Parisi splitting functions are reproduced when one of the initial partons retains its energy, i.e. when x_1 or $x_3 \to 1$. If in eq. (48) or (49) $x_1 \to 1$ and $x_3 = 1-z$, then

$$(1-x_1)\frac{d\sigma}{dx_1dx_3} \propto \frac{1}{z}(1+(1-z)^2). \tag{51}$$

In the case of an initial gluon, this is part of two dipoles. The Altarelli-Parisi function is then the sum of the contributions from both dipoles [7]. Thus with $x_3 \to 1$ and $z = 1-x_1$ in eq. (49) or (50), one gets

$$(1-x_1)\frac{d\sigma}{dx_1dx_3} \propto \frac{1}{z}(1+(1-z)^3). \tag{52}$$

Together with the corresponding term from the other dipole (obtained by $z \to 1-z$) one gets

$$\frac{1+(1-z)^3}{z} + \frac{1+z^3}{1-z} = 2\left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right]. \tag{53}$$

b) Orientation. The cross-sections (48) - (50) give only the shape of the event and not the orientation, i.e. it is not specified how the partons recoil. It was shown in [10] that

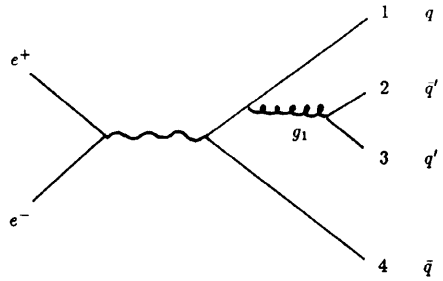

Figure 14: One of the diagrams for $q\bar{q}$ emission.

for an isolated dipole, i.e. for a $q\bar{q}$-system, one gets the correct orientation by defining a probability for one of the quarks to take the entire recoil. If one uses this method for all the dipoles, one does, however, not get satisfactory agreement with the data.

It is reasonable to assume that for a dipole in the middle of the string, i.e. for a gg dipole, the connection with neighbouring dipoles would prevent the gluons from taking large recoils, and that the recoils are distributed so as to cause a minimal disturbance in the shape of the string. Therefore, when a gg-dipole emits a gluon, the authors choose to orientate the subsystem by minimizing $p_{T1}^2 + p_{T3}^2$ [5].

The remaining problem is then how to treat a qg-dipole. In this case the gluon is connected to the rest of the string, while the quark is a 'free' endpoint. The authors assume here that the quark takes the entire recoil. These choices can be described as a minimization of the change in the direction of the colour flow in and out of the emitting dipole.

A further problem arises when an extended parton (e.g. a diquark) takes part in the emission. Then, additional assumptions are needed to dampen gluon emission. These are described in [9].

These may be the most natural choices, but there are other options included in the program (see KAR(3) below).

c) $q\bar{q}$ emission. The extension in ARIADNE 3 to include $q\bar{q}$ emission uses the cross-section of $e^+e^- \to \gamma \to q\bar{q}q'\bar{q}'$ as calculated in [12]. One of the possible diagrams is shown in Fig. 14.

In the limit of low p_T of the gluon g_1 and small invariant mass s_{23} of the emitted $q'\bar{q}'$ pair, one obtains after averaging over the angle of the emitted $q'\bar{q}'$ pair relative to the original $qg_1\bar{q}$ plane:

$$\frac{d\sigma}{dx_1 dx_4 ds_{23} dz} \propto \frac{(x_1^2 + x_4^2)}{(1-x_1)(1-x_4)} \frac{1}{s_{23}} (z^2 + (1-z)^2). \tag{54}$$

Here, x_1 and x_4 are the energy fractions of particles 1 and 4 and z is the light cone fraction of the gluon momentum which is carried away by quark 3.

This is easily factorizable into one part corresponding to the $q\bar{q}$ (initial) $\to qg_1\bar{q}$ emission and one part corresponding to a splitting of the emitted gluon g_1 into the final $q'\bar{q}'$ pair. As expected in this low p_T limit, the latter part reproduces correctly the Altarelli-Parisi splitting function.

There is, evidently, a competition between the possibility to emit a new gluon by either one of the dipoles generated by the emission of g_1 and to split g_1 into a $q'\bar{q}'$

pair. To compare the two processes, the latter is divided into two equal parts, each corresponding to one of the dipoles. This gives

$$\frac{1}{\sigma}\frac{d\sigma}{dx_1 dx_3} = \frac{n_f \alpha_S}{8\pi} \frac{(1-x_2)^2 + (1-x_3)^2}{(1-x_1)} \tag{55}$$

where n_f is the number of possible flavours, α_S is the strong coupling constant and x_2 (x_3) is the energy fraction of the emitted quark (antiquark), (or reversed depending on the colour flow).

d) Ordering. In all competing stochastic processes there is the question of ordering, which for the cascade means whether emission or splitting should be done 'first'. This is treated by means of Sudakov form factors. In conventional parton cascades (e.g. JETSET) a virtuality is assigned to each parton and the emission which results in partons with largest virtuality is done first. In the dipole picture, the p_T of the emitted gluon is the natural ordering variable. Here, an invariant p_T is defined as

$$p_T^2 = \frac{s_{12}s_{23}}{s_{123}} \tag{56}$$

for a gluon 2 emitted from a dipole between particles 1 and 3.

In order to compare gluon emission (i.e. the production of a new gluon) and gluon splitting (i.e the production of a $q'\bar{q}'$ pair), the same ordering variable should be used. Since particles are in general not massless, the more general formula is

$$p_T^2 = \frac{1}{s_{123}} \left[s_{12} - (m_1+m_2)^2\right] \left[s_{23} - (m_2+m_3)^2\right]. \tag{57}$$

Here, s_{123} is the squared total invariant mass of the dipole.

There are possible alternatives to define p_T (see below), but all of them give approximately similar results.

In the program, a p_T^2 and a rapidity $y = 0.5\ln((1-x_1)/(1-x_3))$ is generated in accordance the proper Sudakov form factors for each dipole, for gluon emission and for each possible flavour of gluon splitting. The different p_T's are then compared and the largest one is picked out in each step. This results either in the splitting of a dipole into two (gluon emission) or the splitting of the chain of dipoles into two independent chains (gluon splitting). The process is repeated until none of the p_T's are above a given cut-off (VAR(3)).

e) The physics options in ARIADNE. There are a number of options available through the common block /AROPTN/ (see below).

The first, KAR(1), deals with phase-space restrictions. In the first version of ARIADNE all dipoles were treated independently, which in in general can result in emissions with a larger p_T than the previous ones. This means that the angular ordering is only approximative and it tends to increase the multiplicity significantly for high energies [5]. It also gives a larger rate of gluon splitting [13]. This option is reached by KAR(1)=2. It is also possible to demand that the emissions always have smaller p_T than the previous ones (KAR(1)=0 (default)) in which case the predictions for higher energies are more in line with conventional cascades (e.g. JETSET). There is also a third option (KAR(1)=1) in which the strong ordering in p_T is kept only for gluon emission, allowing the full phase-space for the gluon splitting process. This gives an increase in the gluon splitting without the increase in multiplicity.

The second option, KAR(2), gives a constant α_S (KAR(2)=0) to be specified by VAR(2). Default is a running α_S (KAR(2)=1).

There is no obvious way in which one should treat recoils in the colour dipole approximation. The general strategy in ARIADNE is to minimize the disturbance of the colour flow in neighboring dipoles. If e.g. a purely gluonic dipole between gluon g_1 and g_3, radiates a gluon g_2, the recoils of 1 and 3 are distributed so that p_{T1}^2 and p_{T3}^2 are minimized. If we, on the other hand, have a dipole between a quark and a gluon, there is no further colour flow to disturb on the quark side and the quark is allowed to recoil freely, leaving the gluon in its original direction (KAR(3)=1). There is also an option (KAR(3)=0) for which the quark and gluon recoils are treated in the same way.

The recoil situation becomes more difficult when dealing with extended partons. In this case the extended parton always retains its direction, possibly by emitting an additional gluon to take the recoil (see refs. [2,9] for details). With KAR(4)=0 the special treatment of extended partons is turned off completely.

Apart from the phase-space restrictions in KAR(1) there are two options which directly affect the gluon splitting process. The first is KAR(5) which simply sets the maximum number of flavours allowed in the process.

The other specifies the definition of the invariant p_T used for the gluon splitting to compare it with the gluon emission process. The default option (KAR(6)=0) is to use the generalized version of the p_T defined for gluon emission above (eq. 57). The only difference between gluon emission and gluon splitting is that in the former $m_2 = 0$. This will, however, lead to that a gluon never is split into a $q'\overline{q}'$ pair lying at rest relative to each other. This possibility can be included by using (KAR(6)=1)

$$p_T^2 = \frac{1}{s_{123}} \left[s_{12} - (m_1 + m_2)^2 \right] s_{23}, \tag{58}$$

where partons 2 and 3 are the produced q' and $\overline{q}'$. There is also a third option (KAR(6)=2) with

$$p_T^2 = \frac{1}{s_{123}} s_{12} s_{23} \tag{59}$$

used mainly to study of the effects of different definitions of p_T. It turns out [13] that the rate of gluon splitting is slightly increased with KAR(6)=1 and 2 as compared to KAR(6)=0. The difference is, however, always below 20%.

Fragmentation:
JETSET 6.3 has to be used with ARIADNE 3.0, JETSET 7.1 with ARIADNE 3.1.

Decays: see JETSET

Event Analysis Routines: see JETSET

4.1.4 Comparison with Data and JETSET

In [5] the Monte Carlo results are compared with experimental data on e^+e^- annihilation, and the agreement is found to be very good. At higher energies (LEP and CLIC) some differences are found (e.g. in rapidity and in sphericity distributions) with respect to other parton cascades, so it is left for those accelerators to discriminate.

The most distinct difference is an increased central rapidity plateau (Fig. 15). The relative increase is largest in events with low sphericity. In a 'conventional' (coherent)

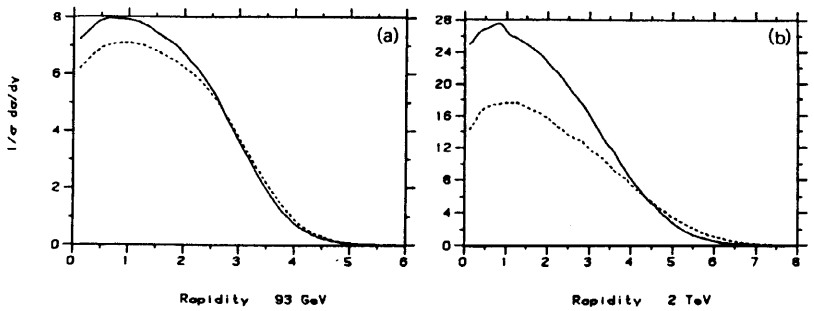

Figure 15: Rapidity distribution at 93 GeV and 2 TeV for the dipole cascade (full line) and for the conventional (coherent) parton cascade of JETSET 6.3 (broken line). In both cases the Lund string fragmentation model is used [5].

parton cascade, the system is easily split into two parts which move rapidly apart. No gluons are emitted centrally. In the dipole scheme there is always one dipole which connects the two parts and can emit gluons in the central region. As a consequence, also the sphericity distribution (Fig. 16) is strongly reduced at low sphericities. Because of increased central emission, the average charge multiplicity changes from $\langle n \rangle$=21.1 and 64.5 at 93 and 2000 GeV in JETSET 6.3 to $\langle n \rangle$=22.6 and 89.4 in ARIADNE version 1 .

Also in ARIADNE 2 with p_T–ordering the agreement with data is very good. However, the predictions for higher energies change drastically. In particular, the multiplicity of gluons increases less rapidly with increasing energy. This might be expected since the phase-space for the gluons gets reduced when demanding that they be softer than the previously emitted gluons. Thus a small change in the treatment of gluons close to the kinematic boundaries will grow in importance when the energy is increased [13].

The major difference between ARIADNE 3 and the JETSET parton shower, is the ordering of successive emissions and hence the application of Sudakov form factors (p_T^2 ordering vs. virtuality ordering). Although both models have the same behavior in the

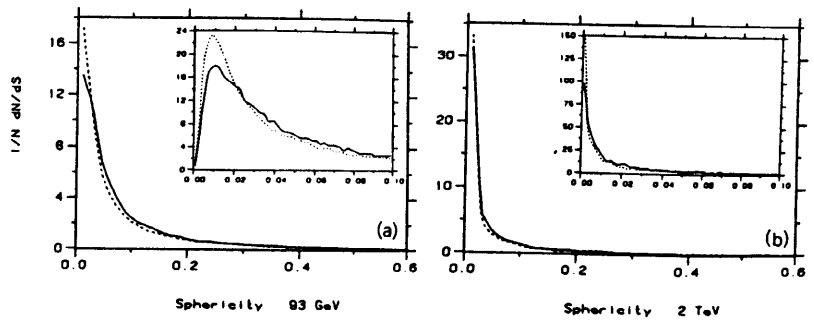

Figure 16: Sphericity distribution at 93 GeV and 2 TeV. Notation as in Fig. 15 [5].

low p_T limit, the difference in ordering makes ARIADNE more likely to emit $q\bar{q}$ pairs. At low energies this effect is smeared out in the fragmentation procedure. At higher energies, however, it results in an excess of heavy quarks in ARIADNE, especially in the low rapidity regions.

4.1.5 Installation and Running

ARIADNE 3.0 is installed at CERNVM (ARIADNE3 FORTRAN on Disk 191 of account KITTEL). It must be loaded together with JETSET 6.3. Running time, including hadronization with JETSET, is 0.03 true IBM 3090 s per 92 GeV event.

ARIADNE consists of a large number of subroutines. Most of these are used internally by the program and are of no real interest for the ordinary user.

In short, ARIADNE works as follows: first the invariant p_T of possible emissions from each dipole is determined (ARDEPT) and the dipole giving the largest p_T is picked out (ARCHDI). This is then allowed to radiate a gluon (ARPLGL) or a $q'\bar{q}'$ pair (ARPLQQ). This is repeated as long as the p_T is larger than the cut-off (VAR(3) in COMMON /AROPTN/).

The main routine in ARIADNE is simply called ARIADNE and is the only routine the normal user has to worry about. Calling ARIADNE causes the program to search the event record (see description of common block /LUJETS/ of JETSET) between entries specified in the /AROPTN/ common block (see below), to find one or more strings of dipoles. These strings will then be made to cascade according to the colour dipole approximation.

The resulting string is ready to be fragmented by a LUEXEC call of JETSET.

To run the program as a black box, all one needs to do is to call SUBROUTINE AREEVT(IFL,ECM). ECM is the center of mass energy of the collision and IFL is the flavour code of the initial quarks to be created. IFL=0 gives a mixture of all allowed flavours according to relevant propabilities. AREEVT generates a complete e^+e^- annihilation event by first generating an initial $q'\bar{q}'$ pair by calling LUXIFL(IFL,ECM,ECM,IFLC), then calling ARIADNE to generate a cascade and finally calling LUEXEC to fragment the resulting strings.

The parameters and switches used in ARIADNE are stored in
COMMON /AROPTN/ IAR(10),KAR(10),VAR(10)

IAR Administrative switches:
IAR(1) (D=1) The first of the particles of the event record considered by ARIADNE.
IAR(2) (D=0) The last of the particles of the event record considered by ARIADNE. If it is set to 0 the last particle is determined by N in /LUJETS/ of JETSET.
IAR(3) (D=0) Maximum number of emissions in ARIADNE. If it is set to 0 there is no limit.
IAR(4) (D=500) Maximum no of particles allowed in the event record in ARIADNE. If this is changed, corresponding changes should be made in COMMON /ARCONF/ and /ARMEMO/.
IAR(5) (D=1) Initialization of the dipole event record. If it is set to 0 the initialization normally made by ARPREP and ARSEPA must be made by the user.
IAR(6) (D=0) Strategy used to set Parameters and switches in ARSEPA:
=0: No parameters are set.

=1: Sets KAR(1)=0, KAR(2)=1, KAR(3)=1,KAR(4)=1, KAR(5)=5, KAR(6)=0, VAR(1)=0.25, VAR(3)=0.50, VAR(4)=1.00, VAR(5)=0.938, PAR(12)=0.35, PAR(31)=0.5, PAR(32)=0.75 (PAR(I) is in the /LUDAT1/ common block of JETSET).
=2: As 1 but KAR(1)=2, KAR(6)=1, VAR(1)=0.10, VAR(3)=0.40, PAR(32)=0.90.
=3: As 1 but KAR(5)=0
=4: As 2 but KAR(5)=0
<0: The same as for IAR(6)> 0, but parameters are only set in the first call.
Note: ARIADNE 3.1 uses PARJ(21), PARJ(41) and PARJ(42) instead of PAR(12), PAR(31) and PAR(32) respectively.
IAR(7) Number of calls to ARIADNE.
IAR(8) (D=1) Fragmentation in AREEVT:
=0: No fragmentation.
=1: The AREEVT call finishes by an LUEXEC call.
IAR(9-10) Not used.

KAR Physical Switches:
KAR(1) (D=0) Phase-space restrictions:
=0: All emitted partons have invariant p_T smaller than the previously emitted partons.
=1: All emitted gluons have p_T smaller than the previously emitted partons. No phase-space restrictions on $q'\bar{q}'$ emissions.
=2: No phase-space restrictions. All dipoles completely independent.
KAR(2) (D=1) α_S:
=0: Constant α_S = VAR(2).
=1: Running $\alpha_S = \alpha_0/\ln(p_T^2/\Lambda_{QCD}^2)$ where p_T^2 is the invariant transverse momentum used as ordering variable.
KAR(3) (D=1) Quark recoils:
=0: All partons treated the same.
=1: Quarks take all recoil if not extended.
KAR(4) (D=1) Extended partons:
=0: No special treatment of extended partons.
=1: Extended partons treated according to ref. [2].
KAR(5) (D=5) Maximum number of flavours in $q'\bar{q}'$ emission.
KAR(6) (D=0) Definition of invariant p_T in $q'\bar{q}'$ emission:
=0: p_T^2 according to eq. (57).
=1: p_T^2 according to eq. (58).
=2: p_T^2 according to eq. (59).
KAR(7) (D=0) Argument in running α_S for gluon splitting
=0: p_T^2 as defined by KAR(6).
=1: $Q^2 = s_{23}$.
KAR(8-10) Not used.

VAR Physics parameters:
VAR(1) (D=0.250) Λ_{QCD} used for α_S.
VAR(2) (D=0.200) Constant α_S (KAR(2)=0).
VAR(3) (D=0.500) Cutoff in invariant p_T.
VAR(4) (D=1.000) Power in soft supression.

VAR(5) (D=0.938) Soft supression parameter.
VAR(6-10) Not used.

References

[1] Ulf Pettersson, 'ARIADNE, A Monte Carlo for QCD Cascades in the Colour Dipole Formulation', LU TP 88-5 (April 1988)

[2] Leif Lönnblad and Ulf Pettersson, 'ARIADNE 2, A Monte Carlo for QCD Cascades in the Colour Dipole Formulation, AN UPDATE', LU TP 88-15 (Sept. 1988)

[3] Leif Lönnblad, 'A Manual for ARIADNE Version 3', LU TP 89-10 (June 1989)

[4] G. Gustafson, Phys. Lett. **B175** (1986) 453

[5] G. Gustafson, U. Pettersson, Nucl. Phys. **B306** (1988) 746

[6] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Phys. Lett. **B165** (1985) 147

[7] B. Andersson, G. Gustafson, Z. Physik **C3** (1980) 223;
B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rep. **97** (1983) 31

[8] H.B. Nielsen, P. Olesen, Nucl. Phys. **B61** (1973) 45

[9] B. Andersson, G. Gustafson, L. Lönnblad, U. Pettersson, 'Coherence Effects in Deep Inelastic Scattering', LU TP 88-14, to be published in Z. Physik **C**

[10] R. Kleiss, Phys. Lett. **B180** (1986) 400

[11] B. Andersson, P. Dahlqvist, G. Gustafson, 'An Infrared Stable Multiplicity Measure on QCD States', LU TP 88-10 (1988)

[12] R.K. Ellis, D.A. Ross, A.E. Terrano, Nucl.Phys. **B178** (1981) 421

[13] B. Andersson, G. Gustafson, U. Pettersson, in preparation

4.2 CALTECH-II

4.2.1 Basic Facts

Program name: Caltech-II [1,2,3]
Version: Caltech-II , from June 1987

Authors: Thomas Gottschalk
CALTECH
Pasadena, CA 91125, USA
Duncan Morris
Physics Department
University of Wisconsin
1150 University Avenue
Madison, WI 53706, USA
BITNET MORRIS@WISCPHEN

Program size: 9800 lines

4.2.2 Physics Introduction

The basic premise of the Caltech-II model [2,3,1] is that the production of a hadronic final state in e^+e^- annihilation can be factorized [4] into three independent phases.

1. The formation of a multi-parton system using the leading log approximation of perturbative QCD, with some modifications to account for gluon coherence effects.
2. The mapping of the parton system onto relativistic strings, and the subsequent evolution of these strings to form low mass colorless *clusters*.
3. The decay of the low mass clusters to hadrons through an elaborate parameterization of low energy data.

Phase 1 of the model rests on a solid theoretical treatment within perturbative QCD, and as such should not really be considered as a model. Phase 2 uses the relativistic string as a model of the QCD flux tube, with confinement and color screening modeled by allowing the strings to break. This phase of the model reflects the current theoretical thinking on confinement, motivated by semiclassical calculations and QCD calculations performed on lattices. Phase 3 is nothing more than a parameterization of existing low energy experimental data. The assumption is that this data is valid in this context, i.e., that factorization is valid, and that low mass clusters produced in low energy experiments hadronize through the same mechanism as low mass clusters produced in high energy experiments.

4.2.3 Detailed Physics Description

Parton showers:
Phase 1 of the Caltech-II model consists of the generation of a multi-parton system. An initial quark antiquark pair is produced with a flavor determined by the quark charge. The first gluon radiated from the $q\bar{q}$ system is weighted by the exact $\mathcal{O}(\alpha_S)$ $e^+e^- \to q\bar{q}g$ matrix element, subsequent branchings in the parton shower are generated according to Leading Log QCD.

The generation of a parton shower follows the prescription outlined in section 2.2.2,

with only minor modifications. For a general branching

$$q^*(p_1) \to q(p_3) + g(p_4), \tag{60}$$

the splitting variable z is taken as

$$z = \frac{1}{2}(1 - t/t_p)(1 + \cos\theta^*) \tag{61}$$

where θ^* is the angle between p_3 and the q^* direction (in the overall CM frame), with the angle evaluated in the rest frame of p_1. t_p is the mass squared of the parent of p_1, and $t = p_1^2$.

In practice the splitting kernels of section 2.2.2 become singular at $z = 0$ or 1, and it is necessary to introduce the constraint that only splittings that produce daughter partons with invariant mass above some cut are allowed. This invariant mass cut t_c then controls the point at which the parton shower evolution is terminated and the second phase of the model is entered.

Color coherence in LLA showers. The incorporation of some next-to leading order terms into the LLA formalism turns out to be particularly simple. This is because it can be shown that the gluon coherence effect is equivalent, after azimuthal averaging, to ordering the emission angles of subsequent radiation down the shower, so that, at any late time in the shower development, the emission angle is smaller than any previous emission angle within the same branch.

In the Caltech-II model by default successive branches in the parton shower are ordered in angle [5,6,7].

String evolution in Caltech-II:
At the end of the development of the parton shower, the parton system is still in a high Q^2 regime, well above the hadron mass spectrum. In phase 2 of the model the partons are mapped onto the ends of strings, that are used to model the behaviour of QCD flux tubes as the system's Q^2 is degraded towards the hadron mass scale. Color screening and its role in the fragmentation process is modeled by allowing the strings to break [8,9,10,11,12].

The partons are mapped onto the ends of the color strings in such a way that each string is a colorless object containing a quark, an antiquark, and some number of gluon kinks. In the leading log QCD shower of phase 1 of the model, information concerning the color flow down the parton shower is maintained, so that the assignment of partons to strings is well defined and unique, this is illustrated in Figure 17.

In Caltech-II the full 3+1 dimensional equations of motion for a massless string are used to calculate the time evolution of the string system [13]. For massless partons the initial momenta serve to completely specify the boundary conditions for the string equations of motion. For massive quarks the formalism is extended, and a small unphysical string section is introduced to take into account the mass of the quark.

Color screening is implemented by allowing the strings to break, i.e., introducing pair creation points in the strings as the strings evolve. String evolution is parameterized in terms of the invariant area (the string analog of proper time for a point particle) swept out in space-time by a string. The invariant area swept out by an open string during one complete cycle of its periodic motion is a function of κ, the (constant) string

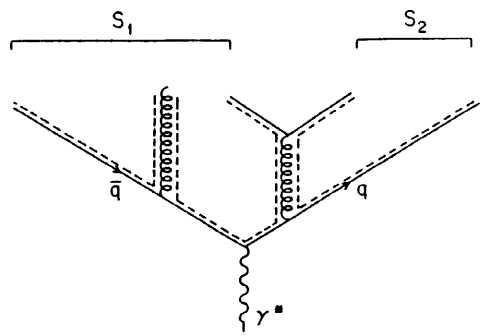

Figure 17: The mapping of a five-parton state onto strings. The color flow in the parton shower is shown as dashed lines.

tension, and the invariant mass, M, of the quark-antiquark pair that define the ends of the string. Since κ and M are the only invariants in the system, the area

$$A = \frac{M^2}{\kappa^2} \tag{62}$$

swept out by the string as it evolves, is also an invariant. With the identification of invariant area as the correct extension of proper time for string-like objects, the radioactive decay law can be generalized to give the probability of a string break occuring per unit area swept out by the string,

$$dP_b = P_0 dA, \tag{63}$$

where P_0 is a constant.

The only constraint put on the choice of a string-break point in Caltech-II is that it should never leave a piece of string so light that it could not be identified with a hadron or at least decay into hadrons —otherwise the break can occur *anywhere* along the string. When the string-break point is chosen too close to the end of the string, the break point is shifted, so as to allow direct string breaks of the form

$$string \to string + hadron. \tag{64}$$

Breaks such as this are controlled by a special parameter in the model, W_{min}, that gives the maximum mass allowed for a string to undergo such a decay.

With the above proviso on string breaks near to the string ends, Caltech-II evolves the initial string system according to an iterative scheme based on the fundamental iteration:

$$string \to string_1 + string_2. \tag{65}$$

The string evolution is terminated by a cut-off procedure, implemented as follows. A parameter W_{max} is introduced in the model, and for the particular string segment being considered, the two particle threshold W_{th} is determined. The quantity W_{cut} determined by

$$W_{cut} = W_{max} + W_{th} \tag{66}$$

is evaluated. If the piece of string remaining after the break has a mass, W, below W_{cut} (i.e., less than W_{max} above threshold), then the string is passed to the cluster

decay routines. If, on the other hand, the mass of the string segment is above W_{cut}, the probability, P_{string}, that it is further evolved as a string (as opposed to being passed to the cluster decay) is taken to be:

$$P_{string} = 1 - \exp\left\{\frac{1}{2}\rho_c(W - W_{th} - W_{max})^2\right\}, \tag{67}$$

where ρ_c is given by

$$\rho_c = P_0/\kappa^2. \tag{68}$$

The quantity ρ_c is a parameter that controls the string dynamics through the string tension and the uniform probability for the occurrence of string breaks. The functional form of eq. 67 is chosen to provide a smooth transition between the string evolution phase and the low mass cluster decay phase of the model.

It is important to note that all the transverse momentum of the final state hadrons comes from the isotropic decay of clusters in the final phase of the model: no attempt is made to add transverse momenta at the string break points.

Generation of Hadrons from Low-Mass Strings:
At the end of the string evolution phase of the Caltech-II model, the low mass string segments are identified with colorless clusters and allowed to decay into hadrons. The association of a cluster with a hadron is made through an elaborate parameterization [2].

Each cluster is fully specified by its mass W and its flavor. The transformation of a cluster (C) into hadrons (H) takes place through a series of two-body decays of the form

$$C \rightarrow X_1 + H_2, \tag{69}$$

where X_1 can be either another cluster or a hadron.

For every possible cluster a list of possible decays is generated with their relative decay probabilities. The decay mode used by the program is then picked from this list at random. The relative decay probability is given by the product of three factors,

$$P = P_f \times P_s \times P_k, \tag{70}$$

where P_f is a flavor factor, P_s a spin factor, and P_k is a kinematic factor.

P_f, the flavor factor. All clusters coming from the string evolution phase of the model are meson clusters i.e., they contain a quark and an antiquark. When such a cluster decays it may produce a meson cluster or a baryon cluster (see Figure 18) depending on whether a quark or a diquark pair are created. In the decay of a baryon cluster, only quark pairs may be selected. In the decay of a meson cluster the new flavour is chosen from the following

$$u, d, s, uu, ud, us, dd, ds, ss, \tag{71}$$

with P_f the weighting factor for each choice.

P_s, the spin factor. For the case of a cluster decay into hadrons, P_s is just the spin multiplicity factor of the final state hadrons,

$$P_s = (2J_1 + 1)(2J_2 + 1), \tag{72}$$

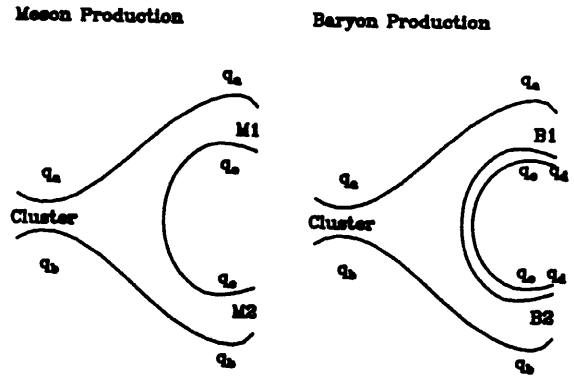

Figure 18: The formation of baryons, mesons, baryon clusters and meson clusters from cluster decays.

where J_1 and J_2 are the spins of the final state hadrons. For decays with clusters in the final state, a meson cluster is given spin 0, and a baryon cluster is given spin 1/2.

P_k, the kinematic factor. P_k is the kinematic suppression factor, proportional to the phase-space available for each decay mode of the cluster. In the case of two hadrons in the final state (of mass M_1 and M_2),

$$P_k = \frac{1}{W^2}[W^4 + M_1^4 + M_2^4 - 2W^2M_1^2 - 2W^2M_2^2 - 2M_1^2M_2^2]^{1/2} = \frac{2}{W}p^*(W, M_1, M_2). \tag{73}$$

When there is a cluster in the final state P_k must be modified. The approximate result

$$P_k = \frac{2}{W}\int_{M_0}^{W-M_1} dM p^*(W, M, M_1)\rho(M) \tag{74}$$

is used. Here $\rho(M)$ is the assumed distribution of cluster masses,

$$\rho(M) = A(M - M_0)^N \Theta(M - M_0), \tag{75}$$

with M_0 a constant depending on the cluster flavor. The parameters A and N are optimized by comparing the model to low energy data.

4.2.4 The parameters of Caltech-II

Although the three phases of Caltech-II appear to be independent of each other, there is considerable overlap between them. The end stages of the parton shower, could (and should) really be considered as the start of the color screening process, that is modeled in phase 2. In a similar way the transition from the string model to the cluster decay phase of the model is not sharp, but rather smooth (as indicated by equation 67). This smoothness, in the transition from one phase of the model to the next, is correct from the physics point of view, but it does mean that the parameter values are highly correlated with each other, and that the parameter tuning is complicated.

By far the largest number of parameters in the model are concerned with the parameterization of the cluster decays. The philosophy of the Caltech-II approach, that hadronization is a universal and local phenomenon, dictate that these parameters be

regarded as fixed and not subject to change after they have been shown to describe the low energy data.

With the provisos above it is possible to identify *five variable* physics parameters in the model.

Parameters that control the parton shower:
There are two parameters that control the generation of the parton shower. The first has already been discussed, this is t_c, which determines when the shower evolution should be terminated. The second parameter is the LLA scale parameter Λ_{LLA}, which enters in the α_S expression. This parameter should not be confused with the true QCD scale parameter, and it should be realised that there is no simple relation between Λ_{LLA} and α_S. While the value of Λ_{LLA} is left free for the user to fix, there does exist a prescription for removing the effect of t_c; t_c should simply be low enough that variation in its value does not effect the predictions of the model. For sufficiently small values of t_c, the effect of additional soft gluons at the end of the parton shower is washed out by the string fragmentation of the parton system.

Λ_{LLA} should be adjusted so that the model predictions for the hard perturbative features of the events are in agreement with the data. A recommended value for Λ_{LLA} would be 600 MeV, with $t_c = 1$ GeV^2.

Parameters that control string evolution:
There are three parameters that control the string evolution phase of the model, ρ_c, the string breaking parameter, W_{max}, the string termination parameter, and W_{min}, the minimum mass of a string segment for normal iteration through string breaking. Of these three parameters, the first two are vastly more important than the third.

As mentioned above all the transverse momenta in the hadron final state comes from the cluster decays. The transverse structure of the events is then completely determined by ρ_c and W_{max}. With W_{max} the reason for this is clear: by increasing W_{max} the mean mass of the clusters is increased and so the transverse momenta are higher, also after decay products are boosted from the cluster center of mass frame.

ρ_c influences the transverse structure of the events by controlling where the first string break is likely to occur. If ρ_c is reduced the strings break less frequently, and on average at a later point in their evolution, where they are longer, and so produce higher mass clusters at the input to phase 3 of the model.

By controlling the mechanism of the string breaks near the ends of the strings, W_{min} is able to influence the distribution of high z particles that eventually appear in the final state. In some sense this is a technical issue, since the same effect could be obtained in this model by adding fictitious masses to the quarks. A nominal value for W_{min} would be 0.25 MeV.

4.2.5 Comparison with Data

The Caltech-II model has been compared to data over the energy range 14 GeV to 46 GeV, by the model authors and in studies made by the Mark II and Mark-J experiments [3,14,15,16].

The Mark II study found that the Caltech-II model was unable to correctly describe the Minor and Thrust distributions at CM energy 29 GeV. The Mark-J study found

good agreement between Caltech-II and the data, over a wide range of CM energies, in all distributions with the exceptions of Thrust and the energy-energy correlation function.

The failure of the Caltech-II model to correctly predict the Thrust distribution and the energy-energy correlation may reflect the absence of $\mathcal{O}(\alpha_S^2)$ corrections in the parton generation, and the inadequacy of imparting transverse momentum to the hadrons through cluster decay alone.

References

[1] T.D. Gottschalk, D.A. Morris, Nucl. Phys. **B288** (1987) 729

[2] T.D. Gottschalk, Nucl. Phys. **B239** (1984) 325

[3] T.D. Gottschalk, Nucl. Phys. **B239** (1984) 349

[4] R.K. Ellis *et. al.*, Phys. Lett. **B78** (1978) 281

[5] T.D. Gottschalk, *in* ed. H.-U. Bengtsson, C. Buchanan, T.D. Gottschalk, A. Soni, Observable Standard Model Physics at the SSC: Monte Carlo Simulations and Detector Capabilities (World Scientific, Singapore, 1986), p. 122

[6] Yu.L. Dokshitzer *et. al.*, 'Coherence Effects in QCD Jets', Leningrad Nuclear Physics Research Institute report 1051 (1985)

[7] G.C. Fox, S. Wolfram, Nucl. Phys. **B168** (1980) 285

[8] D. Amati, G. Veneziano, Phys. Lett. **B83** (1979) 87

[9] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rep. **97** (1983) 33

[10] X. Artru, G. Mennessier, Nucl. Phys. **B70** (1974) 93

[11] X. Artru, Phys. Rep. **97** (1983) 147

[12] A. Casher *et. al.*, Phys. Rev. **D20** (1979) 179

[13] D.A. Morris, Nucl. Phys. **B288** (1987) 717

[14] D.A. Morris, 'The relativistic string in the Caltech-II model of hadronization and electron-positron annihilation', Ph.D. thesis, CALT-68-1440 (1987)

[15] H. Stone, 'Experimental tests of QCD and fragmentation in e^+e^- annihilation', Ph.D. thesis, California Institute of Technology (1988)

[16] Mark II Collaboration, A. Petersen *et. al.*, Phys. Rev. **D37** (1988) 1

4.3 COJETS

4.3.1 Basic Facts

Program name: COJETS
Version: COJETS 5.12 from May 1989
Author: Roberto Odorico
Department of Physics
University of Bologna
Via Irnerio 46
I-40126 Bologna, Italy
BITNET M6ZBOX51 @ ICINECA2
Decnet CINE88::M6ZBOX51
Tel: +39-51-242018
Fax: +39-51-247244
Program size: 22126 lines (including documentation)

4.3.2 Physics Introduction

COJETS originated as a Monte Carlo generator for $e^+e^- \to$ hadrons. Its first quantitative results were presented at the ECFA-LEP Workshop held in Rome, September 1979 [1,2]. It later evolved as a generator for pp and $\bar{p}p$ collisions at large p_T [3]. There are short term plans of reintroducing e^+e^- annihilation in it. Its updated documentation can be found in [4] and in the program file.

For $\bar{p}p$ and pp collisions, the interaction process is first calculated at the parton level by means of perturbative QCD and the electroweak standard model. Partons are then independently fragmented into jets of hadrons, according to the Field-Feynman model, and the beam jets contribution is added, according to a longitudinal phase-space model. Multiple QCD radiation off initial and final partons is included. The following processes are simulated:

1. purely hadronic production, including particles with charm, bottom and top;
2. weak Drell-Yan production of the W and Z vector bosons; and
3. Drell-Yan production of the charged lepton pair continuum.

In leptonic decays of the W and Z bosons multiple QED radiation effects are considered. Decays of top particles are treated for an arbitrary top quark mass, with multiple QCD radiation effects for decay jets and multiple QED radiation for decay leptons included.

Plans for the new version of e^+e^- annihilation are based on usage of the routines for parton shower development, parton fragmentation into hadrons and Z decay already existing in the program, which have been thoroughly tested and compared with data. The simulation would include:

1. γ/Z^0 cross-section with dependence on initial and final polarizations;
2. multiple QED radiation off the beams;
3. multiple QED radiation in leptonic final states;
4. top decay of whatever mass; and
5. multiple QED and QCD radiation in top semileptonic decays.

4.3.3 Detailed Physics Description

The following presentation refers only to plans for e^+e^- annihilation. A detailed documentation for pp and $\bar{p}p$ interactions can be found in [4] and in the program file.

Electroweak cross-section:
The γ/Z^0 cross-section includes dependence on beam and final lepton polarizations.

Initial state QED radiation:
Multiple initial state radiation is included, treated in the leading log approximation (LLA). The main aim is to calculate the energy loss undergone by beam particles.

Relative flavour composition:
It can be generated by the program, as an option.

Perturbative QCD approach:
The final state at the parton level is calculated by an LLA (timelike) parton shower algorithm [1]. The algorithm reproduces analytic results within statistically accuracy, once kinematic constraints which are neglected in analytic calculations are disregarded. Using an appropriate choice of the gauge, the jets from the initial quark and antiquark are generated as a single shower, thus essentially eliminating the uncertainty associated with the choice of initial evolution scales for the two primary jets. The parton shower algorithm is corrected so as to reproduce the $\mathcal{O}(\alpha_S)$ results for single gluon emission, with the same procedure which is used for QED radiation in $Z^0 \to l^+l^-$ [5].

Final state QED radiation:
When leptons are produced at the primary vertex, multiple QED radiation is computed in the LLA, corrected so as to reproduce $\mathcal{O}(\alpha_{em})$ results for single photon emission in the final state [5]. This treatment is meant to be appropriate for muons and taus, and essentially focuses on collinear radiation. Appropriate interfacing with other matrix element programs for leptonic final states is under consideration.

Fragmentation:
Final partons are converted into particles by the Field-Feynman independent fragmentation model, duly extended to heavy flavour quarks and baryons. A corrective procedure ensures global conservation of energy and momentum.

Decays:
Unstable particle decay modes are specified in a table, which can be changed by the user. By default, decay distributions are treated according to phase-space, except for specific decays (like, e.g, semileptonic decays of heavy flavour particles), where matrix elements are known. Top particle decays include multiple QCD and QED radiation effects (see [6], where a comparison with the Jezabek-Kühn $\mathcal{O}(\alpha_S)$ calculation for semileptonic decays can also be found). The top quark mass can be arbitrary.

4.3.4 Comparison with Data

A comparison of the early version for e^+e^- annihilation with data available at that time can be found in [2].

Extensive comparisons have been done with $\bar{p}p$ data. A comparison with UA1 data at large transverse energy can be found in [7].

4.3.5 Installation and Running

COJETS is written in FORTRAN 77. It is maintained with the PATCHY code management system, which is part of the CERN Program Library. It is supported for CRAY, IBM, VAX and APOLLO machines.

Besides the native user interface, also an ISAJET type interface is available. If a consensus is reached for a common standard of the output format, optional generation of the output according to the standard will be implemented.

4.3.6 Future Plans and Availability

COJETS is available from the CERN Program Library and on request from the author. An updated e^+e^-version will appear shortly.

References

[1] R. Odorico, Nucl. Phys. **B172** (1980) 157, *erratum* **B178** (1981) 545, Computer Physics Commun. **24** (1981) 73

[2] P. Mazzanti, R. Odorico, Phys. Lett. **B95** (1980) 133, Z. Physik **C7** (1980) 61

[3] R. Odorico, Nucl. Phys. **B228** (1983) 381, Phys. Rev. **D31** (1985) 49, Computer Physics Commun. **32** (1984) 139, *erratum* **34** (1985) 431, Computer Physics Commun. **32** (1984) 173, *erratum* **34** (1985) 437

[4] R. Odorico, University of Bologna preprint DFUB 88-27 (1988)

[5] S. Laporta, R. Odorico, Nucl. Phys. **B266** (1986) 633, Computer Physics Commun. **39** (1986) 127

[6] R. Odorico, University of Bologna preprint DFUB 89-2 (1989)

[7] UA1 Collaboration, C. Albajar *et. al.*, Z. Physik **C36** (1987) 33

4.4 DPSJET

4.4.1 Basic Facts

Program name: DPSJET, a Statistical Model of Jet Evolution

Author: Wolfgang Ochs
Max-Planck-Institut für Physik und Astrophysik
– Werner-Heisenberg-Institut für Physik –
P.O. Box. 40 12 12,
Föhringer Ring 6
D-8000 München 40, Fed. Rep. Germany
BITNET WWO @ DM0MPI11

Program size: 481 lines

4.4.2 Physics Introduction

Most present day jet models are based on a partonic branching process as derived from QCD. Though there cannot be much doubt about the correctness of QCD at short distances, the models involve some arbitrariness and a considerable complexity for the hadronization phase at the end of the cascade. Some time ago the author developed a rather simple model, which tries to derive the main features of the data from a statistical equipartition principle [1], following in spirit the model of Hagedorn [2], but extending towards the regime of short distances. While limited in its applicability, as we shall see, it can serve as an interesting toy model for some properties.

4.4.3 Detailed Physics Description

Basic Assumptions:
An important concept of the model is the 'dynamical phase-space': energy and momentum are quantized in a volume which expands with velocity of light in a sequence of discrete time steps. The system evolves in this phase-space in a stochastic manner, and it is assumed that any quantum state resolvable at a time can be occupied with the same probability. This hypothesis of 'instantaneous equipartition' determines uniquely the distribution over the available states and leads to a branching process. Neglecting all degrees of freedom except energy and momentum and restricting to final state pions a minimal model emerges with no other parameters than h, c and m_π (the length scale which enters Hagedorns model as volume parameter is now determined by the light velocity c after a given evolution time). There are though some theoretical approximations in the elaboration of this scheme (rectangular volume, independent evolution of different branches and others).

Predictions:
Strength of coupling. This is predicted by the equipartition principle, a unique feature of the model not shared by other models known to us. The model reproduces well the shape and magnitude of the angular energy-energy correlation (except for a region of small angles), which is often used to determine the coupling parameter α_S of QCD. It also predicts the particle multiplicity which comes out, however, too low by

$\sim 30\%$. Agreement with data can be obtained by choosing a smaller (unphysical) pion mass.

Scale invariant branching. Heavier states of mass M decay into 2 body states with masses m_1, m_2 with a scaling distribution $F(\mu_1, \mu_2)$, $\mu_i = m_i/M$ until the pion mass is reached. This is similar to QCD branching, but the effective α_S is not running in the above scheme. This yields a multiplicity rise like $\overline{n} \sim W^{0.44}$, KNO multiplicity scaling, scalebreaking of inclusive spectra, and predictions for energy flows and jet phenomena [1].

Hadronic intermediate states. The intermediate states carry hadronic quantum numbers, so there is no recombination phase at the end of the cascade and the only scale in the model is the mass of the final hadron. Consequently the scaling laws, KNO scaling for example, set in already early, close to threshold [3]. Also strong 'intermittency' effects are predicted [4]. Though a branching of hadronic states looks like a rather unconventional concept, QCD after confinement could yield effectively such a structure ('exclusive parton hadron duality' [5]).

Conclusion:
The model represents a simple realisation of a scale invariant hadronic branching process. Its theoretical strength and at the same time its phenomenological weakness is the lack of adjustable parameters. (One could of course adjust the function $F(\mu_1, \mu_2)$ to the data). It reproduces well the main trends of data, though not always at a quantitative level. Also, because of the limitation in the degrees of freedom included (only pions, no resonances, no spin), some more specific questions cannot be adressed. The model may be of practical use today for the derivation of scaling laws and size estimates of any desired observable, so that tests of the general structure of scale invariant hadronic branching processes can be obtained.

The strongest support for the basic equipartition ansatz we see in the correct prediction of the coupling strength. At the moment this scheme looks rather different from a gauge tieory such as QCD. A similarity can be seen in the fact, that both the gauge principle and the principle of instantaneous equipartition in dynamical phase-space determine completely the interaction for a given set of symmetries (conservation laws), the latter even the coupling strength. It remains to be seen whether in a more complete treatment a closer connection between both schemes can be established.

4.4.4 Installation and Running

The program may be obtained from the author. It is also found in the file DPSJET FORTRAN on the TORSJO 192 disk on CERNVM. The program contains comments for what to change in order to run at different energies, etc.

References

[1] W. Ochs, Z. Physik **C23** (1984) 131, Z. Physik **C15** (1982) 227

[2] R. Hagedorn, Nuov. Cim. Suppl. **3** (1965) 147

[3] W. Ochs, Z. Physik **C34** (1987) 397

[4] W. Ochs, J. Wosiek, Phys. Lett. **B214** (1988) 617

[5] G. Cohen-Tannoudji, W.Ochs, Z. Physik **C39** (1988) 513

4.5 EPOS

4.5.1 Basic Facts

Program name: EPOS, a Monte Carlo Program for e^+e^- Annihilation using the Fire-String Model [1]
Version: EPOS 2.0 from June 1989
Authors: †L. Angelini, L. Nitti, †M. Pellicoro
I.N.F.N. and †Department of Physics - Bari (Italy)
G. Preparata
I.N.F.N. and Department of Physics - Milano (Italy)
G. Valenti
I.N.F.N. - Bologna (Italy)
BITNET GIANNI @ VXCERN.CERN.CH (G. Valenti)
Program size: 5400 lines

4.5.2 Physics Introduction

The EPOS authors start from a small number of well established and universally accepted basic notions of hadro-dynamics. They are:

1. the hadrons are extended objects made of point-like constituents ('quarks') with spin 1/2 and fractional charges;
2. the forces between quarks obey the colour SU(3) symmetry realized through a gauge principle;
3. the quarks are confined inside the hadronic matter where they move 'quasi-freely' at short distances.

QCD is accepted as an elegant and appealing scheme, embodying the above-stated principles, but it is noted that the QCD computational scheme is not yet powerful enough to describe properly the quark confinement property. The most popular theoretical models make use of the 'quasi-free' behaviour and calculate the quark interactions as free objects. The connection between quarks and hadrons, however, are treated outside the QCD scheme by means of fragmentation models, and this obscures the predictive strength of QCD in its comparison with experiments.

The authors therefore depart from the traditional perturbative picture of QCD. Instead the theoretical framework proposed by G. Preparata is used, Quark-Geometro-Dynamics (QGD) [2] and Anisotropic Chromo-Dynamics (ACD) [3]. In this approach, gluons are not considered as dynamical entities, but only as carriers of a confinement force. The only physical states allowed in the model are colour singlets made out of quarks. More specifically:

- hadrons are made by point-like spin-1/2 'quasi-free' quarks, confined in limited space-time domains, whose dimensions increase linearly with the mass of the hadrons;
- there is a basic perturbative structure in the number of quarks that take part in a given reaction;
- the hadrons emerge in the final state through a sequential decay of a small number of 'Fire-Strings' (FS), highly excited quark-antiquark states.

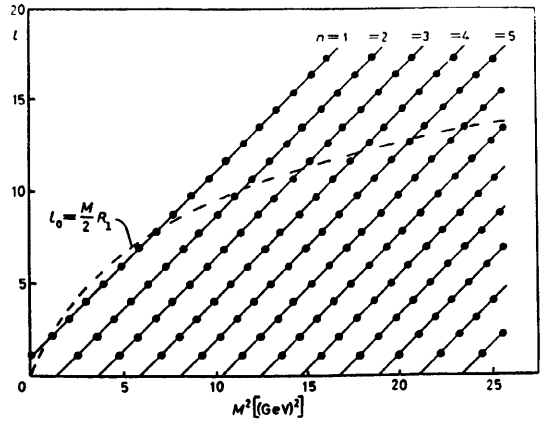

Figure 19: The Chew-Frautschi plot of the $q\bar{q}$ states in QGD.

The physical reasons given for these principles are the following:

- the success of the quark model indicates that the confining forces become effective only at large distances, where the behaviour is well described by linear potentials;
- the Zweig rule strongly suggests the existence of an underlying perturbative structure in the dynamics of hadrons;
- as a consequence of the FS dynamics, a high degree of universality of quark fragmentation, i.e the jet-like structure, can be found in the final states, not only in e^+e^- annihilations, but also in lepton-hadron and hadron-hadron collisions.

4.5.3 Detailed Physics Description

The Fire-String:
The spectra of the quark-antiquark and three quark states have been constructed with the following inputs:

1. m_q, the quark masses ($q = u, d, s, c, b$);
2. R^2, the rate of increase of hadronic radii with hadronic masses (thus $R^2 \propto 1/\kappa$, κ being the ordinary string tension).

Then one obtains a set of Regge trajectories for the quark-antiquark states (Fig. 19) well described asymptotically by the formula:

$$M^2_{n,l} \simeq \frac{\pi}{R^2}(2n + l), \tag{76}$$

where l is the angular momentum and n is the radial excitation quantum number. Calling $p_{1,2} = p/2 \pm k$, the four-momenta of the quark and the antiquark, respectively, one can also derive the excited meson wave-functions, whose asymptotic structure (M^2 large) is given by:

$$\psi_{n,l}(p, k) \simeq F(p_1^2, p_2^2) Y_l^m(\Omega_k), \tag{77}$$

where Ω_k is the direction of the relative $q\bar{q}$ momentum k in the meson rest frame. $F(p_1^2, p_2^2)$ is a universal function, which exhibits approximate factorization:

$$F(p_1^2, p_2^2) = G(p_1^2)G(p_2^2). \tag{78}$$

The quark 'propagator function' $G(p_1^2)$ has the following behaviour:

$$G(p^2) \sim \frac{\sin(R^2(p^2 - m_q^2))}{p^2 - m_q^2}, \quad p^2 \approx m_q^2$$
$$G(p^2) \sim \frac{1}{p^2}, \quad |p^2| \gg m_q^2, \tag{79}$$

and correctly represents confinement (i.e., no singularities on the mass shell and free field behaviour at short distances).

The quasi-degeneracy in the mass $M_{n,l}$ of states with different angular momenta (see Fig. 19) leads to the construction of FS states, whose wavefunctions are:

$$\psi(p,k;\Omega) = \sum_{l=0}^{l=l_0} \sum_{m=-l}^{m=+l} \psi_{n,l,m}(p,k) Y_l^m(\Omega) \tag{80}$$

with $p^2 = M^2$ held fixed and Ω an arbitrary spatial direction. The maximum value of the angular momentum l_0 in eq. (80), given by

$$l_0 = R_T M/2 \tag{81}$$

with $R_T \simeq 0.5$ fm, decouples states of higher angular momenta in order to maintain a finite range (or at most logarithmically increasing with energy) to the strong interaction.

In the configuration space the structure of the state of eq. (80) is depicted in Fig. 20, comprising 'quasi-free' quark and antiquark plane waves moving with three-momentum $|p| \simeq M/2$ along the direction Ω and confined in domains with longitudinal dimensions increasing linearly with the mass M of the state. The transverse radius R_T, on the other hand, remains fixed and does not depend on the mass of the state. Thus the origin of the concept 'Fire-String' of mass M and axis along Ω. Its structure is intended to reproduce, as closely as allowed by quark confinement, the free states of the parton model.

The FS is therefore a universal hadronic structure: it gets produced in a variety of ways depending on the production process, but it should leave its footprints in the structure of final states by means of its subsequent decays. This gives rise to the high degree of universality claimed in the high energy reactions.

Production mechanism of Fire-Strings:
In order to describe the physics of the electron-positron annihilation at the center of mass energy W, the process

$$e^+e^- \to \gamma/Z^0 \to FS \tag{82}$$

is considered, where the virtual γ/Z^0 of mass W couples to a FS of mass W.

As a visualization of this process, Fig. 21 shows a $q\bar{q}$ pair, comprising the initial FS, generated in a direction θ with respect to the incident direction. The quark flavour composition (u, d, s, c and b quarks are included) and the angular distribution are in agreement with standard electroweak predictions, see section 2.1.

Mesonic decays of a Fire-String:
Once a FS of mass M is produced according to some particular mechanism, the problem of describing and calculating the FS-decay into stable hadrons follows the principles introduced and motivated in the preceding sections.

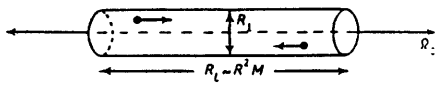

Figure 20: The spatial structure of the state ψ in eq. (80).

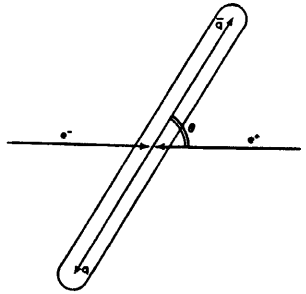

Figure 21: The production of the initial FS in e^+e^- annihilation.

In the fluctuating colour field of the FS, a $q\bar{q}$ pair is created, which induces a break-up of the colour flux with the creation of two FS's of smaller masses (see Fig. 22). The process of pair creation being a slow one, it is quite unlikely that two or more pairs get created simultaneously; thus the basic decay mechanism for a FS turns out to be:

$$FS \to FS_1 + FS_2, \tag{83}$$

followed by similar processes involving FS_1 and FS_2, until all the produced FS's have masses of the order of hadronic low-lying states (π, ρ, K, K^*, etc.). As shown in Ref. [4], one finds that the amplitude of the process (83) has the important property that it scales, and can be expressed in terms of the scaled variables:

$$x_{1,2} = \left(\frac{M_{1,2}}{M}\right)^2,$$
$$\chi = \left(\frac{p_T}{M}\right)^2, \tag{84}$$

where M, M_1, M_2 are the masses of FS, FS_1, FS_2, respectively, and p_T is the decay momentum transverse to the orientation of the initial FS.

In particular, note that the scaling property in p_T/M implies that large transverse momenta may be generated in the decay of a heavy FS. It is this mechanism which leads to the emergence of multijet phenomena in EPOS, not perturbative QCD emission of gluons, as in most other programs. Further comments can be made:

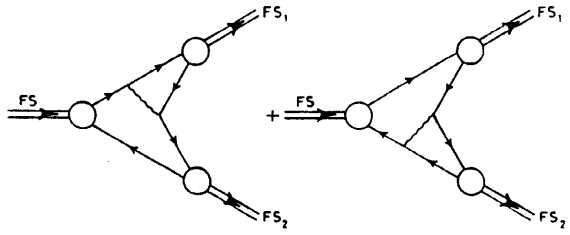

Figure 22: The basic mechanism for the mesonic FS-decay.

- one has demonstrated the suppression of quantum interference effects, thus making it possible to calculate the FS-decay process in terms of probabilities rather than quantum amplitudes;
- the mechanism of $q\overline{q}$ creation strongly suppresses heavy quark pairs, thus, to a good approximation, only light quark pairs need to be considered (heavy flavours occur only when the initial FS contains a heavy quark).

For a detailed discussion concerning the kind of physics at work in FS-decay, consult the review paper [4].

Baryonic decays of a Fire-String:
When a di-quark pair $qq - \overline{qq}$ gets created in the FS colour field, the FS decays into a baryon-antibaryon pair (see Fig. 23)

$$FS \rightarrow B_1 + \overline{B}_2, \tag{85}$$

where B_1 ($\overline{B}_2$) is a lowest-lying baryon (antibaryon) belonging to the SU(6) 56-plet. The amplitude has been computed in Ref. [5], where details can be found.

The calculational framework is not powerful enough to determine the relative normalization between the mesonic and baryonic FS-decay probabilities, so that it must be introduced as an input to be fitted by experiments. This input parameter is the only new parameter entering the description of FS-decay, all the other parameters (m_q's, R^2) having been fixed by the structure of the mass spectrum of the low-lying hadrons.

The full FS-decay process is now completely determined and it gives rise to the 'tree-like' picture shown in Fig. 24.

4.5.4 Comparisons with Data

The theoretical model sketched in the previous sections has been extensively applied to the analysis of electron-positron annihilation into hadrons at high energies ($W > 10$ GeV). The results, reported in Refs. [4,5,6,7], have been obtained by the Monte Carlo program EPOS [8]. It is worthwhile to mention that this program has only one tunable input parameter, i.e. the relative normalization between the baryonic and the mesonic decay modes of a FS of a given mass.

4.5.5 Installation and Running

The program is installed at CERN; contact G. Valenti (BITNET address given in the header) for further details. A recent, updated manual exists [1], with sufficient information to get going. In addition, a somewhat older manual [8] gives much further information, most of which is still relevant.

References

[1] L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, BARI TH/89-47 (1989)

[2] G. Preparata, Phys. Rev. **D7** (1973) 2973;
for a broad review see G. Preparata, Quark-Geometrodynamics: A New Approach

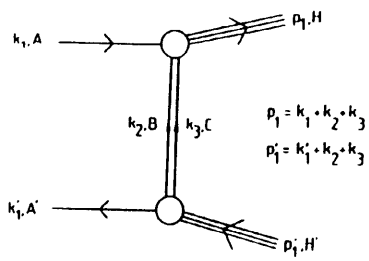

Figure 23: The basic mechanism for the baryonic FS-decay.

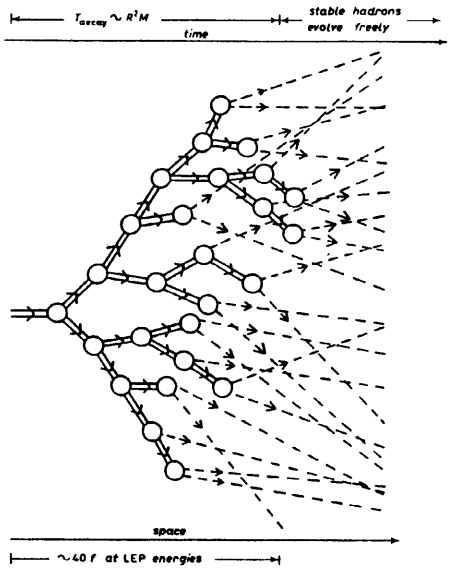

Figure 24: The typical tree-like structure of a full decay evolution of a FS.

to Hadrons and Their Interactions, in: The Why's of Subnuclear Physics, ed. A. Zichichi (Plenum Press,1979) pp. 727-803

[3] For a review of ACD, see G. Preparata, Quark Confinement and Lepton Liberation in an Anisotropic Space Time, *in* Fundamental Interactions (Cargèse 1981), eds. M. Levy, J.L. Basdevant, D. Speiser, J. Weyers, M. Jacob and R. Gastmans (Plenum Press, 1982) pp. 421-448

[4] L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, Riv. N. Cim. **6** (1983) 1

[5] L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, Phys. Lett. **B123** (1983) 246

[6] L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, Phys. Lett. **B119** (1982) 456.

[7] L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, Phys. Rev. **D27** (1983) 1668.

[8] L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, Comp. Phys. Comm. **34** (1985) 371.

4.6 EURODEC

4.6.1 Basic Facts

Program name: EURODEC, a Monte Carlo program for the fragmentation of partons and decay of particles [1,2,3]

Version: EURODEC 2.3, from May 1989

Authors: A. Ali
Department of Theoretical Physics
DESY
Notkestrasse 85
D-2000 Hamburg 52, FRG
B. van Eijk
EP Division
CERN
CH-1211 Geneva 23, Switzerland

Program size: ~ 4600 lines, plus ~ 2200 lines particle (decay) table

4.6.2 Physics Introduction

In the years 1983 and 1984, the analysis of the rapidly increasing amount of hadron collider data from experiments at the CERN Sp$\bar{p}$S collider showed the need for a dedicated Monte Carlo program for the simulation of QCD and weak hard scattering processes, going beyond lowest order in perturbation theory. This resulted in a special purpose Monte Carlo program called EUROJET [1], written with the main emphasis on heavy flavour production in hadron-hadron collisions. The EUROJET derivatives 'EUROPPB' [2] and 'EURODEC' [3] have recently been made available as relatively independent packages, with a clear separation in functionality. The EURODEC package, subject of this chapter, signals the start of renewed activity of the authors to match up with the need for programs with a broad range of applications, i.e. also suitable for detailed studies in e^+e^- interactions. The start-up of LEP, and later HERA, were therefore the main motivations for separating the parton production code (as for hadron physics in EUROPPB) from the fragmentation and decay part (EURODEC). Both packages may, in principle, easily be interfaced with other programs.

4.6.3 Detailed Physics Description

Parton fragmentation:
Due to the application of the program in conjunction with higher order matrix elements, only very little effort has, up to now, been put in describing parton fragmentation in terms of a parton shower model. The independent parton fragmentation scheme in EURODEC has nevertheless evolved with time, and some improvements with respect to the 'old' Field-Feynman [4] picture have been introduced. As the IF model has been described in detail in section 2.3.1, we limit ourselves to the discussion of the main extensions that have led to the Modified Independent Fragmentation (MIF) model.

Following the hierarchy in the FF cascade procedure, the fraction of energy and

longitudinal momentum carried by the hadron is defined as

$$z = \frac{E^H + p_{\parallel}^H}{E^Q + p_{\parallel}^Q}. \tag{86}$$

FF (and others) have proposed probability distributions for z. For the relatively light quarks (u, d, s) the mass of the initial quark will hardly have any effect on the fragmentation procedure, and one may use

$$f(z)dz = 1 - a + 3a(1-z)^2. \tag{87}$$

At each stage of the fragmentation process the production of a diquark-antidiquark pair is considered, and occurs with a fixed probability. Since there is only very little information available on the fragmentation of diquarks and antidiquarks, one assumes that they fragment in a similar way as do single quarks. For the fragmentation of light diquarks (limited to combinations of u, d and s quarks) one has assumed the parametrization of eq. (87). For heavier quarks (c, b, t, ...) the fragmentation function a la Peterson *et. al.* (see eq. (28)) [5] has been introduced. This function peaks closer to $z = 1$ as m_Q increases, reflecting the hard fragmentation function of heavy flavours.

At each fragmentation step, both the newly formed hadron and the parton receive a relative transverse momentum. The $\bar{p}_T$ of the hadron balances the $\bar{p}_T$ of the 'left-over' parton, which acts as the starting point for the next step in the cascade. Since the $\bar{p}_T$ is balanced at each step, the transverse momentum sum of all particles (including the final left-over parton) equals the transverse momentum of the initial quark. The $\bar{p}_T$ may be different at each step of the fragmentation, and is generated according to the probability distribution

$$f(p_T)dp_T^2 \propto \exp\left(-\frac{(2-z)p_T^2}{4z(1-z)\sigma^2}\right) dp_T^2. \tag{88}$$

The function expresses the phase-space suppression in the regions where either $z \to 1$ (all parent quark energy is transfered to the hadron) or $z \to 0$ (minimum energy transfer to the hadron). The individual momentum components which build up the transverse momentum are generated with a random orientation in space.

Flavour is conserved at each step of the cascade. Thus, after the treatment of n fragmenting partons, we are left with a set of n jets and n soft partons. Except for the overall violation of energy-momentum conservation, due to the introduction of massive particles in the fragmentation chain (recall that, in in general, $m_Q < m_H + m_q$), flavour and transverse momentum remain conserved quantities.

Gluons are treated as light quark-antiquark pair jets of random flavour. The energy of a quark Q is determined via the relation $E_Q = xE_{gluon}$, with for x the probability distribution (the $g \to Q\overline{Q}$ splitting function [6]):

$$f(x)dx = \frac{3}{2}\left[x^2 + (1-x)^2\right] dx. \tag{89}$$

Consequently, the antiquark receives the fraction $1-x$ of the gluon energy. Both Q and $\overline{Q}$ are given a relative p_T, eq. (88) and are fragmented independently using the modified FF scheme. The width of the p_T distribution is assumed to be similar for both quark and gluon fragmentation.

After all partons have been fragmented, the left-over partons are combined in the original rest frame of the n parton system, with the help of an algorithm that orders all possible allowed (soft partons emerging from the same gluon are not allowed to re-combine) parton combinations according to their separation in ΔR ($\Delta R = f(\phi, \eta)$, where η is pseudorapidity). Finally the energy and momentum sum for the hadrons is determined. Since energy/momentum conservation is violated in the fragmentation process, the total energy after fragmentation (E^{sum}) will in general not be equal to the total centre-of-mass energy (Q_0) one started with. In addition, one has $\overline{p}^{sum} \neq \overline{0}$. The hadrons are therefore Lorentz transformed to the frame where $\overline{p}^{sum} = \overline{0}$. The boost vector for this transformation is defined by E^{sum} and $\overline{p}^{sum}$. The mass of the partonic system is invariant under Lorentz transformations. The discrepancy in energy remains, and can be minimized by rescaling the momentum $\overline{p}_H$ of each individual hadron by a factor α,

$$\Delta Q = \left(\sum_{i=1}^{n} \sqrt{(m_H^i)^2 + \alpha^2 |p_H^i|^2} \right) - Q_0, \tag{90}$$

where n is the total number of hadrons. The correct α is determined (iteratively) by the requirement that $\Delta Q = 0$, and is typically close to unity [7].

It should be stressed that the kinematics involved in the decay of particles is not affected by the energy-momentum conservation procedure sketched in this section. All decays are taken care of after fragmentation and after energy-momentum conservation is imposed.

Particle Decays:

Both heavy flavour and non-heavy flavour decays have been extensively studied by the authors. The particle decay table contains a very extensive list of decay modes, as reported by several experimental groups, and relies heavily on Particle Data Group information [8]. For only a very limited number of particles, 100 % of the decay channels have experimentally been resolved. The bulk of decay modes (baryons and heavy flavours) remain unobserved and have to be modelled in order to obtain a consistent Monte Carlo treatment. A huge list of particles and particle decay modes ($\sim$ 2000 lines) has been set up, in which particle mass, lifetime and type of decay (in terms of angular correlations among particles etc.) have been defined. The user can have direct access to this table with the help of any local editor. Therefore, modifications of and/or extensions to the particle list and/or particle decay table are a simple exercise, and are meant for those users that want to include their own model or a specific decay mode, without going into the details (code) of the program.

In this section, we briefly discuss and summarize the decay matrix elements as included in the present version of the program (2.3). We start by describing some typical non-heavy flavour decays. For example, the kinematics of the decay products in the decay $\omega, \phi \to \pi^+\pi^-\pi^0$ is well described [9] by convoluting the pure phase-space with

$$|M|^2 \propto |\overline{p}_{\pi^+} \times \overline{p}_{\pi^-}|^2. \tag{91}$$

Similarly, the complete kinematics of the Dalitz decays of the π^0 and η ($\pi^0, \eta \to \gamma e^+e^-$) is correctly described by applying the 'Kroll-Wada procedure' [10], in which one generates a properly distributed virtual mass for the e^+e^- pair. The energy and momentum of the electron and positron are obtained via the pure phase-space decay of the virtual object.

Pseudoscalar particles PS_2 and PS_3 produced in the decay chain $PS_0 \to PS_1$ + Vector $\to PS_1 + PS_2 + PS_3$ are constrained to have the following decay angular distribution in the rest frame of the vector :

$$|M|^2 \propto \cos^2\theta_{02}, \tag{92}$$

where θ_{02} is the opening angle between the PS_2 and PS_0 momentum vectors [11].

Tau lepton decays are among another class of thoroughly studied decays. Although the τ puzzle has not yet been solved completely, the amount of data on exclusive decay modes with up to seven charged particles has accumulated rapidly. An extensive discussion on τ decay matrix elements and branching ratios can be found in [12] and references quoted therein.

In the modelling of heavy flavour decays, a pure $V - A$ structure is assumed for the weak charged current in $Q \to q\overline{f}f$. All fermions (Q, q, f and $\overline{f}$) may have arbitrary masses. The normalized matrix element squared reads [13]

$$|M|^2 = C \frac{8\left[(p_q p_f) \cdot ((p_Q + m_Q s) p_{\overline{f}})\right]}{m_Q^4}, \tag{93}$$

with p and m the momentum four-vector and mass of the indicated particles. C contains the propagator and coupling constants, while s describes the polarization (four-vector) of Q (see also section 6.2 and [3]). If we sum over polarizations and treat the decay of Q in its rest frame, eq. (93) reduces to the very simple expression

$$|M|^2 = \frac{16C}{m_Q^4}(p_q p_f)(m_Q E_{\overline{f}}), \tag{94}$$

where $E_{\overline{f}}$ is the energy of the antifermion.

The semileptonic decays $X \to l\nu_l Y$, with X charm or bottom pseudoscalar particles (the strong or electromagnetic decay of charm and bottom vectors will take place at a much shorter time scale than their weak decay), and Y predominantly strange or charm pseudoscalar particles, respectively, are governed by the same matrix element, eq. (93), whereas decays $X \to l\nu_l Y^*$, with Y^* a vector particle are described assuming a pure vector current. The matrix element for this type of decays can be found in ref. [14]. The normalized lepton energy distribution can be written as

$$\frac{1}{\Gamma}\frac{d\Gamma}{dE_l} = \frac{96}{m_X^5 f} \frac{E_l^2(m_X^2 - m_{Y^*}^2 - 2m_X E_l)^2(m_X^2 + 2m_{Y^*}^2 - 2m_X E_l)^2}{(m_X^2 - 2m_X E_l)^2}, \tag{95}$$

with m and E the appropriate masses and energies of the particles, and f the usual phase-space factor.

Although the effect of the W propagator may safely be neglected in case of charm decays, it has to be included in the decay of heavy third and fourth family quarks and/or leptons. The effect becomes extremely important in decays of, for instance, top quarks with mass $m_t \sim m_W$ or larger. If we return to the process $Q \to q\overline{f}f$, and assume that the heavy quark Q approaches or exceeds the physical mass of the W, and $m_Q \gg m_q$, the kinematics of the $f\overline{f}$ pair will be dominated by the W propagator. This is apparent from the Breit-Wigner expression,

$$B.W. \propto \frac{1}{\left[m_W^2 - (p_f + p_{\overline{f}})^2\right]^2 + m_W^2 \Gamma_W^2}, \tag{96}$$

Table 4: Decay matrix element codes. A † in the third column indicates that, depending on the available phase-space, each parton will give rise to a hadronic jet, while ‡ denotes that a pair of partons will give one single hadron.

Matrix element	Type	Partons allowed?
0	up to 6-body phase-space	no
1	$\eta, \pi^0 \to e^+e^-\gamma$	no
2	$\omega, \phi \to \pi^+\pi^-\pi^0$	no
3	$PS_0 \to PS_1 + V \to PS_1 + PS_2 + PS_3$	no
4	$\tau^- \to \nu_\tau K^-, \nu_\tau \pi^-$	no
5	$\tau^- \to \nu_\tau K^{*-}, \nu_\tau \rho^-, \nu_\tau a_1^-$	no
6	$a_1 \to \rho\pi \to \pi\pi\pi$	no
12	onium $\to$ gluon gluon	yes
13	onium $\to$ gluon gluon gluon	yes
14	onium $\to$ gluon gluon γ	yes
15	onium $\to H^0\gamma$	no
22	weak $V-A$ decay	yes †
23	weak Vector decay	yes †
24	onium single quark decay	yes †
30	phase-space	yes ‡
52	weak $V-A$ decay	yes ‡
53	weak Vector decay	yes ‡

with m_W and Γ_W the physical mass and width of the $W^\pm$, and $(p_f + p_{\overline{f}})^2$ the virtual W mass-squared. The upper and lower limits on the virtual mass are fixed by the masses of the particles involved in the decay. The function in eq. (96) will strongly peak if the virtual mass gets close to or exceeds the physical W mass. Ergo, for a heavy top quark or heavy fourth generation particles, the Breit-Wigner will significantly influence the decay distributions.

Obviously, hadronic decays of very heavy flavours will be dominated by a large jet multiplicity. Both f and $\overline{f}$ may give rise to one or more jets of particles. Additional fragmentation of the resulting quark 'q' may also occur. In EURODEC, each of the $f-\overline{f}$ and q–spectator quark/diquark systems are treated as colour neutral objects.

In Table 4 the decay matrix element codes defined for EURODEC 2.3 are shown, together with their application. Due to the introduction of kinematical correlations between particles, the ordering of the decay products in the decay table is fixed for the majority of matrix elements!

By default, the physical particle masses are taken from the particle table as read by the EURODEC program at initialization time. However, the user may select a physical width (optional), which is at present set to twice the full width half maximum ($\sim \pm\Gamma$) of a Breit-Wigner distribution, thus assuming that the narrow width approximation is a sufficiently good approximation for the mass distribution. The Γ values are calculated from the particle lifetimes, which are tabulated in the particle table as well. The 'smeared' mass is then generated according to

$$m_{smear} = \sqrt{m_{phys}\{m_{phys} + \Gamma \tan[A_- + (A_+ - A_-)R]\}},$$

$$A_\pm = \arctan\left[\frac{(m_{phys} \pm \Gamma)^2 - m_{phys}^2}{m_{phys}\Gamma}\right], \tag{97}$$

with R a uniformly distributed random number in the range $0 < R < 1$. The expression for m_{smear} is a straightforward inversion of the analytical integration of the Breit-Wigner formula [1].

In summary, the present version of the program handles the decay of (light and heavy) pseudoscalar mesons, vector mesons, a limited set of tensors etc., spin 1/2 baryons, spin 3/2 baryons, and leptoquarks. Charmonium, bottomonium and toponium decays, although not discussed here, have been modelled in detail. There are in principle no physical limits set on the number of particles that could be included in the particle and particle decay tables. The treatment of more exotic particles may be introduced by the user, or can be made available on request.

4.6.4 Comparison with Data

In the past, the program has been heavily used, and is still being used, in the analysis of hadron collider data. A separate event generator was attached to simulate different QCD and weak hard scattering processes. Comparisons of the fragmentation and decay models used in EURODEC with data can be found in ref. [1] and references quoted therein. Recently, the particle and particle decay tables, as well as the matrix elements for heavy flavour and τ lepton decays have been reviewed and brought up to date [15,12]. A note on the modelling of toponium decays is in preparation [16] and will appear soon (summer 1989).

4.6.5 Installation and Running

The package is written in standard Fortran 77 and is available from the CERN central computer program library, in Patchy format or as pure Fortran source code. The program executes at any machine (including Personal Computers), provided the machine has a standard Fortran compiler. No external (library) routines are needed. EURODEC contains its own random number generator (cycle: 2^{144}), which gives a machine independent sequence of random numbers. Depending on the word length of the machine, a very limited set of calculations may have to be performed in double precision. The Patchy version of the program will take care of this automatically if the proper 'machine flag' is selected before compilation. For 32 bit machines, the following flags have been included: *IBM, *VAX, *APOLLO, *MAC and *GOULD. For machines with a word length > 32 bits, the flags *CDC and *CRAY have been introduced. Examples of cradles as used on a VAX for creation of a library and running of test programs are given below.

Library creation:

```
+EXE.
+USE, *VAX.                    Machine selection
+USE, EUDLIB.                  Patch contains all EURODEC +use cards
+PAM, LUN = 11, T = ATTACH.    EURODEC.PAM
+QUIT.
```

Test program:

```
+EXE.
+USE, *VAX.                          Machine selection
+USE, EURODEC.                       Patch contains all EURODEC keeps
+USE, TESTF. (or TESTD)              Test program for fragmentation (or decays)
+USE, patch, deck.                   User selected patch, deck (optional)
+ADD, patch, deck, line nr.          Add call to user supplied routine (optional)
      CALL USER
+USE, USER.                          Select user code (optional)
+PAM, LUN = 11, T = ATTACH.          EURODEC.PAM
+PATCH, USER.
+DECK, USER.                         User supplied deck (optional)
      SUBROUTINE USER
      ...
      ...
      RETURN
      END
+QUIT.
```

Both test programs, TESTF and TESTD, have a main routine and should therefore not be selected at the same time. The test examples make use of the CERN histogram package HBOOK, which should be included in the link step.

Before a first event can be treated, a call to the routine EUDINI *has* to be made (see also TESTF and/or TESTD). This routine executes a set of sub-programs (contained in the EURODEC package), and is responsible for the proper initialization of the program. Title cards and particle/particle decay tables are read. Since the authors have refrained from using a BLOCK DATA subprogram, one single subroutine (EUDCIN) performs the pre-setting of parameters in common blocks. An advantage of this method is that one can have a completely machine independent program implementation.

If the EURODEC package has to be incorporated in a larger software environment, the user may prefer to call the routine EUDCIN before EUDINI has been executed. As a result, the subsequent call to EUDINI at a later stage will ignore the routine EUDCIN. This procedure allows the user to modify pre-settings in the common blocks, at the overall initialisation time of his own program.

Similarly to the particle and particle decay tables, EURODEC titles are supplied in a separate (Fortran accessible) file. User modifications can easily be made with the help of any local editor. The titles are characterized by a sequence of keywords, qualifiers and values, generally in this order. A single keyword may be accompanied by more than one qualifier. A list of keywords and qualifiers is included in the user manual as a 'tear-off' page, whereas a model title file is supplied via the CERN program library.

Program entries:

As stated above, at least one call to the initialisation package EUDINI (no calling arguments) has to be made before any other routine is called. The event handling package contains three main entries which have different functionalities. The user may treat them as entries to a 'black box', provided the correct input information is supplied:

CALL FRAGMT(IP1, IP2, NTEIL0)

performs the fragmentation (and subsequent decay)/decay of IP1 - IP2 partons/particles contained in the common /MOMGEN/. The result is returned in the common /RESULT/ at location NTEIL0 + 1 (NTEIL0 serves as an off-set) and onwards. A call to the routine

DECAYS(NTEIL0, NTEIL1)

will result in the full cascade decay of particles at location NTEIL0 to NTEIL1 in the common /RESULT/. Emerging particles are added to the list in /RESULT/. The decay of a single particle, preventing a full subsequent decay cascade can be obtained via a call to

PARDEC(IP, IPPAR, IPDEC),

where IP points to the location at which the decaying particle is located in /RESULT/, IPPAR is a pointer to the entry in the particle table, and IPDEC is the pointer to the first line in the decay table for this specific particle. Explicit examples of applications can again be found in the user manual [3].

4.6.6 Future Plans and Availability

Minor modifications and updates will be included on a regular basis. A major upgrade of the program will be released as EURODEC version 3.0, which will become available at the end of summer 1989. By then the program will include a set of utility routines, a parton shower model (including the branching of both quarks and gluons), and an interface to the Standard HEP Common Block for Monte Carlo event generators (see section 6.2). This new version (including documentation) will be obtainable from the CERN central computer program library (DD Division).

References

[1] See for instance:
A. Ali, B. van Eijk, E. Pietarinen, unpublished;
B. van Eijk, PhD Thesis, University of Amsterdam, unpublished (1987);
A. Ali, B. van Eijk, I. ten Have, Nucl. Phys. **B292** (1987) 1

[2] 'EUROPPB User Manual' and 'EUROPPB Reference Manual', in preparation

[3] B. van Eijk, 'EURODEC User Manual', DELPHI Report 89-39 PHYS 41 PROG 136 and DD Long Write-up, 'EURODEC Reference Manual', DELPHI Report and DD Long Write-up, in preparation

[4] R.D. Field, R.P. Feynman. Phys. Rev. **D15** (1977) 2590, Nucl. Phys. **B136** (1978) 1

[5] C. Peterson, D. Schlatter, I. Schmitt, P. Zerwas, Phys. Rev. **D27** (1983) 105

[6] G. Altarelli, G. Parisi. Nucl. Phys. **B126** (1977) 298

[7] This method was first used in a Monte Carlo program for e^+e^- interactions called QCDJET, see A. Ali, E. Pietarinen, J. Willrodt, DESY Internal Report T-80/01 (1980)

[8] Particle Data Group, G.P. Yost *et. al.*, Phys. Lett. **B204** (1988) 1

[9] R.H. Dalitz, 'Strange Particles and Strong Interactions', London, Oxford University Press (1962)

[10] N. Kroll, I. Wada, Phys. Rev. **98** (1955) 1355

[11] K. Gottfried, J.D. Jackson, Nuovo Cim. **33** (1964) 309;
G.F. Wolters, Proc. Koninklijke Academie van Wetenschappen, **B67** (1964) 192

[12] B. van Eijk, J. Fuster, DELPHI Report 89-45 PHYS 42, see contribution to Working Group on Tau Physics, this Workshop

[13] See for instance: S. Jadach, J.H. Kühn, Max Planck München Report MPI-PAE/PTh 64/86 (1986)

[14] A. Ali, Z. Physik **C25** (1979) 41

[15] K. Bos, B. van Eijk, '(Heavy Flavour) Decays in Eurojet', DELPHI Report 89-50 PHYS 44 (1989) DELPHI, see contribution to Heavy Flavour Working Group, this Workshop

[16] E. Barberio, B. van Eijk, DELPHI Report, in preparation (1989)

4.7 HERWIG

4.7.1 Basic Facts

Program name: HERWIG, Hadron Emission Reactions With Interfering Gluons [1,2]
Version: HERWIG 3.2 from February 1989
Authors:
G. Marchesini
Dipartimento di Fisica
Universita di Parma
INFN, Gruppo Collegato di Parma
Italy
DECnet VAXPR::MARCHESINI
B. R. Webber
Cavendish Laboratory
University of Cambridge
Madingley Road
Cambridge CB3 0HE
United Kingdom
BITNET BRW1@PHY-HEP.CAM.AC.UK
Program size: 6183 lines

4.7.2 Physics Introduction

The HERWIG Monte Carlo is a general purpose event generator for the simulation of either time-like or space-like quark-gluon parton showers. Because of its generality, HERWIG can simulate the QCD showers which develop in proton-proton, proton-antiproton, electron-proton or electron-positron collisions, amongst others, all in a uniform theoretical framework. The parton showers proceed through a sophisticated machinery, which accounts for correlations between the parton 4-momenta due to spin and the coherent emission of soft gluons. The treatment of soft gluon emission is based on the so-called Coherent Parton Branching formalism [3]. Coherent Parton Branching is an extension of the Leading (collinear) Logarithmic Approximation (LLA), which sums the leading large logarithms associated with the collinear (mass) singularities of QCD, to include also the leading infrared logarithms, which arise from soft gluon emission and related virtual corrections. It also takes into account the class of nonleading logarithms associated with gluons that are soft but not collinear, or collinear but not soft. The suitability of the Leading Logarithmic Approximation for the description of higher order QCD final states is discussed in section 2.2 and, in more detail, in the report of the QCD theory group.

HERWIG has traditionally been more complete with regard to its simulation of QCD interference phenomena than have been the other principal QCD generators, e.g. JETSET. At the time of this writing, HERWIG offers the only parton shower simulation of the jet calculus type which includes azimuthal correlations between partons due to coherent gluon emission, either within a jet or between different jets (ARIADNE also possesses this feature, but is not of the 'jet calculus' type). Note that by 'jet' we mean the quark-gluon shower initiated by a parton that participates in a point-like scattering process. HERWIG is also unique, at the time of this writing, in that it includes the full machinery to deal with coherent soft gluon emission in space-like parton showers.

This latter is relevant e.g. for partons which evolve off-shell inside a proton and then participate in a hard subprocess in a proton-antiproton collision.

Here, the discussion of the HERWIG parton shower algorithm will focus exclusively on its application to electron-positron collisions, for which partons begin their evolution with a large virtuality and subsequently evolve to lower virtuality (rather than the other way around). The virtual masses-squared of the partons are always positive, and it is the time-like shower algorithm alone which is of relevance.

An important aspect of the time-like parton shower algorithm of HERWIG is that interference phenomena within the shower due to the parton spins are included to full leading collinear (LLA) order. HERWIG is again unique amongst shower simulations, at the time of this writing, in possessing this feature.

HERWIG also contains a model for hadronization. This hadronization model is of the 'cluster algorithm' type first introduced by Wolfram [4]. Wolfram type models take advantage of the preconfinement characteristic [5] of perturbative QCD. Preconfinement is the property of QCD whereby a parton almost always finds itself nearby in momentum space (and real space) to a parton carrying opposite color charge, whenever the partons have evolved from high to low virtuality, as in a parton shower. In such a case, to leading order in the number of colors, each color may be naturally associated with its anti-color – at least in principle – in a low mass and necessarily colorless object. This object, generically called a 'cluster,' therefore possesses the properties of a simple hadron resonance.

In the case of gluons, which possess two color indices, this statement remains true independent of how the gluon's momentum is shared between its colour and anticolour. In the Monte Carlo simulation, it is necessary to assume a mechanism whereby the colour indices and associated momentum fractions of a gluon are assigned to distinct clusters. In HERWIG, all gluons present at the end of the perturbative shower are forcibly split into quark-antiquark pairs. This mechanism (the 'Wolfram ansatz') is a simple way to make the necessary colour assignments.

Hadrons are (usually) produced in HERWIG through two-body, isotropic decays of the clusters. The hadron species in these two-body decays are selected in proportion to a phase-space factor, which includes the spin degeneracy of the hadronic final state. Thus, in the decay of a $u\bar{u}$ cluster, a (K^+, K^-) state is less likely to be produced than is a (π^+, π^-) one, simply because it weighs more. Similarly, the production of a $(f_2(1270), \pi^0)$ state is favoured over that of a $(a_1^0(1260), \pi^0)$ state in the decay of a $u\bar{u}$ cluster (of enough mass), because of the spin factor. Thus the spectrum of hadrons is controlled by the invariant mass scale of the clusters rather than by the values of additional phenomenological parameters. In this sense, HERWIG can be said to possess a 'dynamic' mechanism for the creation of hadrons. Similarly, the transverse momentum distribution of hadrons is limited, as required to describe jets, as a natural consequence of the low cluster mass scale, thus ultimately as a consequence of the preconfinement property of perturbative QCD as implemented within the parton shower phase.

The HERWIG Monte Carlo may therefore be characterized as possessing a very detailed treatment of parton shower development, combined with a simple, intuitive model for hadronization, which represents in practice a realization of 'parton-hadron duality' [6], i.e. the postulate that the distribution of hadrons as observed in an experiment is directly related to the underlying distribution of partons.

4.7.3 Detailed Physics Description

Electroweak cross-section and selection of flavour:
HERWIG does not contain any algorithms for electromagnetic radiation, and thus no direct photons are generated from either the initial or final state. The $q\bar{q}$ system created in the process $e^+e^- \to q\bar{q}$ therefore always has a mass value equal to the center of mass energy E_{CM}, as selected by the user.

When operated as a quark-pair generator for e^+e^- annihilations, the relative probability for producing a $q_f\bar{q}_f$ pair of flavor f is calculated using the Born level expression

$$\frac{d\sigma}{d\Omega}(e^+e^- \to q_f\bar{q}_f) = \frac{\alpha_{em}^2}{4s} \cdot \left[A_f(1+\cos^2\theta) + B_f\cos\theta\right], \tag{98}$$

with the definitions

$$\begin{aligned}
A_f &= N_{color}\left\{Q_f^2 + 2Q_eQ_fv_ev_f \cdot \frac{s-m_Z^2}{(s-m_Z^2)^2+(m_Z\Gamma_Z)^2} \cdot \frac{s}{4\sin^2\theta_W(1-\sin^2\theta_W)}\right. \\
&\quad \left. +(v_e^2+a_e^2)(v_f^2+a_f^2) \cdot \frac{1}{(s-m_Z^2)^2+(m_Z\Gamma_Z)^2} \cdot \left(\frac{s}{4\sin^2\theta_W(1-\sin^2\theta_W)}\right)^2\right\}, \\
B_f &= 4N_{color}\left\{Q_eQ_fa_ea_f \cdot \frac{s-m_Z^2}{(s-m_Z^2)^2+(m_Z\Gamma_Z)^2} \cdot \frac{s}{4\sin^2\theta_W(1-\sin^2\theta_W)}\right. \\
&\quad \left. +2v_ev_fa_ea_f \cdot \frac{1}{(s-m_Z^2)^2+(m_Z\Gamma_Z)^2} \cdot \left(\frac{s}{4\sin^2\theta_W(1-\sin^2\theta_W)}\right)^2\right\}, \\
v_f &= (I_3)_f - 2Q_f\sin^2\theta_W, \\
a_f &= (I_3)_f, \text{ i.e. } 3^{rd} \text{ component of weak isospin,} \\
Q_f &= \text{electric charge.}
\end{aligned} \tag{99}$$

Thus there are no corrections to account for the different quark masses, nor is an energy dependence included for the width Γ_Z or for the coupling constant α_{em}. Therefore the relative probability for producing the quark pair $q_f\bar{q}_f$ is given by a step function: it equals the magnitude of the coefficient A_f in (98) if the energy E_{CM} is above the threshold $2m_{q_f}$, while it is zero below threshold.

Once the quark flavor has been selected, the polar angle orientation of the back-to-back quark and antiquark is generated randomly according to the distribution in square brackets in equation (98). The azimuthal angle is generated randomly, assuming a uniform distribution between 0 and 2π. The quarks q_f and $\bar{q}_f$ are each given an energy equal to $E_{CM}/2$ and a momentum assuming that they are on-shell.

Parton shower algorithm, overview:
The parton shower simulation in HERWIG requires, as input, the back-to-back on-shell quarks whose creation was discussed in the previous section (we remind our readers that the discussion here will focus exclusively on the parton showers of e^+e^- annihilations). This parton shower simulation is accomplished in three stages. In the first stage, the parton branchings are generated, namely the number and types of partons as well as the mother-daughter relationships. In the second stage the exact kinematics is specified, i.e. a 4-momentum is assigned to each parton. It is at this second stage that azimuthal correlations between the partons within a shower are defined (due to coherence and spin).

Further, gluons are split (non-perturbatively) into quark-antiquark pairs in preparation for hadronization. In the third stage, the overall e^+e^- annihilation event is assembled.

Please note in the description below that the back-to-back q and $\overline{q}$ created in the electroweak scattering are denoted q_0 and $\overline{q}_0$, respectively, in order to emphasize that they are the initial partons of the event. Please note also that by 'jet' we mean the ensemble of partons created in the evolution of either q_0 or $\overline{q}_0$, including the initiating parton q_0 or $\overline{q}_0$ itself. With this rather strict convention, all events of HERWIG are 'two-jet events' by definition.

An option of the HERWIG Monte Carlo permits the generation of '$e^+e^- \to q\overline{q}g$' events (with a user-selected maximum thrust of the $q\overline{q}g$ system), thus resulting in the production of three-jet events, but of three-jet events *only*. In this option the fragmentation and interference of the three jets are handled in the same basic manner as for the two-jet final states, except that one jet is initiated by a gluon. Although we shall mainly concentrate our discussion on two-jet systems, we shall mention below where the three-jet option possesses important differences relative to the two-jet case.

Parton shower branchings (first stage):
Sudakov form factors. To determine whether a parton will branch or not, HERWIG employs the formalism of Sudakov form factors. Roughly speaking, the Sudakov form factor $\Delta_i(Q^2, Q_0^2)$ gives the probability that an off-shell parton i of virtuality Q will evolve to a lower virtuality Q_0 without the emission of resolvable radiation, i.e. that a parton will reach its mass shell (or the infrared cut-off for massless partons) without experiencing any branching whatsoever. Formally, a Sudakov form factor embodies the leading virtual corrections to the same order as those for real branching, and may be expressed as the exponential of an integral over the usual QCD splitting probabilities (see [1] for more details):

$$\Delta_i(Q^2, Q_0^2) = \exp\left[-\int_{Q_0^2}^{Q^2} \frac{dQ'^2}{Q'^2} \sum_j \int dz \frac{\alpha_S(Q'^2, z)}{2\pi} P_{i\to jk}(z)\right]. \tag{100}$$

The functions $P_{i\to jk}(z)$ are the Altarelli-Parisi splitting kernels. Notice that the scale at which the running coupling α_S is evaluated depends on both the virtuality Q'^2 and the branching momentum fraction z. In fact, in order to treat the limiting values $z \to 0$ and 1 in a manner which is considered optimal, HERWIG makes the choice $z(1-z)Q'^2 \simeq p_T^2$, i.e. the transverse momentum-squared of the branching, for the scale of of α_S [7].

Because of the factorization permitted by the exponential form (100), the probability for branching as well as for non-branching may be expressed using Sudakov factors. Thus the probability $B_i(Q^2, Q_{br})$ that parton i will evolve from scale Q to a lower scale Q_{br} without branching may be written [1]:

$$B_i(Q^2, Q_{br}^2) = \frac{\Delta_i(Q^2, Q_0^2)}{\Delta_i(Q_{br}^2, Q_0^2)}. \tag{101}$$

Since the derivative of this function with respect to Q_{br} must represent the differential distribution of Q_{br} at the first branching, starting at virtuality Q, the Monte Carlo simulation of the distribution of Q_{br} is obtained simply by solving the equation $B_i(Q^2, Q_{br}^2) = R$, where R is a uniform random number between 0 and 1. Thus the relation (101), with the modification to be discussed in the next few paragraphs, is the governing equation for deciding whether a parton will branch in HERWIG.

The HERWIG Evolution Variable. A most important aspect of the HERWIG shower algorithm is its incorporation of soft gluon interference phenomena, namely the incorporation of calculations which account approximately for the coherent nature of soft gluon emissions. The Coherent Parton Branching algorithm used in HERWIG is based on the fact that the main result of coherence is a destructive interference effect, which suppresses branching outside an 'angular-ordered' region of phase-space [8].

In HERWIG, this angular ordering is introduced by substituting a new evolution variable ζ in place of the parton virtual masses Q of equation (101). For a parton branching $i \to j + k$, the evolution variable ζ_i is defined by

$$\zeta_i = E_i\sqrt{\xi_{jk}}, \tag{102}$$

$$\xi_{jk} = \frac{p_j p_k}{E_j E_k}, \tag{103}$$

where p_j and p_k are the daughter 4-momenta and E_j and E_k are their respective energies. The variable E_i is the energy of the parent parton i. Thus the HERWIG evolution variable ζ refers in part to the parton *branchings* (through the variable ξ_{jk}) and in part to the individual partons themselves (through the variable E_i). For daughter partons whose energies are large compared to their virtual masses, ξ_{jk} is approximately equal to $1 - \cos\theta_{jk}$, where θ_{jk} is the angle between j and k in the branching. As a consequence, the angular ordering of successive branchings is approximately equivalent to ordering of ξ_{jk}.

A value for the evolution variable ζ may be assigned to an individual parton as follows. Let ζ_i be the value of the evolution variable defined for the splitting $i \to jk$ (equation (102)). The evolution variables of the daughters j and k then have the initial values

$$\begin{aligned}(\zeta_j)_{initial} &= z\zeta_i, \\ (\zeta_k)_{initial} &= (1-z)\zeta_i,\end{aligned} \tag{104}$$

where z is the fraction of the parent energy E_i acquired by parton j. The assignments (104) follow from conservation of the energy E_i in (102). Because of the ordering of the angular variable ξ, these initial values are the upper limits on the evolution variables of i and j at subsequent branchings. The actual values are chosen according to the appropriate Sudakov form factors (expressed in terms of ζ rather than the virtuality Q), as explained above. If j and k do branch, initial values for the evolution variables of *their* daughters may be assigned through the analogous relations to equations (102) and (104).

In the description below, the value ζ for the evolution variable at a particular branching is generically called 'ζ_{br}', while the value ζ for the evolution variable below which radiation can no longer be resolved is generically called 'ζ_0.'

Through the substitutions $Q \to \zeta$, $Q_0 \to \zeta_0$ and $Q_{br} \to \zeta_{br}$ in equation (101), HERWIG incorporates the (leading infrared) interference due to the coherent emission of soft gluons *inside* a jet, in so far as it is equivalent to an ordering of polar emission angles of the partons. The interference *between* jets is taken into account in the same approximation by ordering the angles of initial emissions in each jet with respect to the angle between the two interfering hard partons.

Soft gluon coherence also leads to correlations between the azimuthal emission angles of the partons. Outside the region of ordered polar angles, the azimuthal angles are implicitly averaged over to give no net emission. Inside the ordered region, HERWIG takes account of azimuthal correlations due to the interference associated with coherent gluon emission, by a method that will be presented later in this section.

The first branching. For the generation of the parton shower, the quark and antiquark created in the electroweak scattering process are evolved separately, i.e. each initiates a separate jet of partons, which are combined at a later stage. The algorithm for the evolution of q_0 and of $\overline{q}_0$ is the same; therefore only the evolution of q_0 will be described here.

As a first step, the initial value for the evolution variable ζ must be assigned to q_0, to set the correct upper limit on ζ at any subsequent branching of this parton. Applying the procedure implied by equations (102) and (104), the angular variable $\xi_{q_0\overline{q}_0}$ for the initial event branching $\gamma/Z^0 \to q_0\overline{q}_0$ has the value

$$\xi_{q_0\overline{q}_0} = \frac{p_{q_0}p_{\overline{q}_0}}{E_{q_0}E_{\overline{q}_0}} = 1 - \frac{\vec{p}_{q_0}\cdot\vec{p}_{\overline{q}_0}}{E_{q_0}E_{\overline{q}_0}} \geq 1, \tag{105}$$

so that the initial value ζ_{q_0} for the evolution variable of q_0 is

$$\zeta_{q_0} = z_{q_0}E_{CM}\sqrt{\xi_{q_0\overline{q}_0}} = E_{q_0}\sqrt{\xi_{q_0\overline{q}_0}}. \tag{106}$$

Equation (106) follows from the facts that the parent energy for the initial branching equals the center of mass energy, and that q_0 and $\overline{q}_0$ share this energy evenly, i.e. the energy fraction z_{q_0} in equation (106) equals 0.5. While the value $\xi_{q_0\overline{q}_0}$ in equation (105) is not exactly known (since it depends on the masses the q_0 and $\overline{q}_0$ acquire by the shower evolution), it is larger than unity, since q_0 and $\overline{q}_0$ are back-to-back. As discussed in [1], it is difficult to apply the coherent parton branching formalism consistently when the angular variable ξ is larger than unity. However, the formalism depends only on the quantities $E_i\sqrt{\xi_i} = \zeta_i$ and the energy fractions z_i. Therefore, in the HERWIG algorithm, it is possible to rescale the initial angular variable $\xi_{q_0\overline{q}_0}$ to unity, provided the energy E_{q_0} is correspondingly rescaled to maintain the value of ζ_{q_0} in equation (106):

$$\begin{aligned} \xi_{q_0\overline{q}_0} &\Longrightarrow \xi'_{q_0\overline{q}_0} \equiv 1, \\ E_{q_0} &\Longrightarrow E'_{q_0} = E_{q_0}\cdot\sqrt{\xi_{q_0\overline{q}_0}}, \\ \zeta_{q_0} = E_{q_0}\cdot\sqrt{\xi_{q_0\overline{q}_0}} &\Longrightarrow E'_{q_0}\cdot\sqrt{\xi'_{q_0\overline{q}_0}} = \text{invariant}. \end{aligned} \tag{107}$$

This redefinition therefore involves rescaling the energy of q_0 by the factor $\sqrt{\xi_{q_0\overline{q}_0}} \simeq \sqrt{2}$. This rescaling should in principle have no effect on the parton shower which is subsequently generated, since the initial value of the evolution variable is unchanged (by construction), and since the parton shower development depends on the value of the evolution variable alone. In practice, some small terms (of order $\zeta_0^2/E_{q_0}^2$, where ζ_0 is the cut-off) are introduced when the parton shower from q_0 is boosted back to the original q_0 frame, as part of the final stage of the HERWIG shower simulation; as will be described below. Such 'higher-twist' terms are within the uncertainties of any parton shower simulation and should not constitute a problem at LEP energies. Similar ambiguities, arising essentially from the necessity to use a non-covariant gauge, arise in

the definition of the energy-momentum fraction z, and are common to all parton shower simulations.

The next step in the simulation is to determine if q_0 will branch, i.e. $q_0 \to q + g$, and is accomplished in the following manner.

1. The value of the Sudakov factor $\Delta_{q_0}(\zeta_{q_0}^2, \zeta_0^2)$ is calculated, c.f. equation (101). Here ζ_{q_0} is the initial value of the evolution variable assigned to q_0 (relation (106)), and ζ_0 is the cut-off below which radiation is assumed unresolvable. Since the process being considered is $q \to q + g$, the cut-off ζ_0 is set equal to $m_{q_0} + m_g$, where m_{q_0} is the on-shell mass of q_0, and where m_g is a formal mass value for the gluon which serves as the main perturbative cut-off parameter of the model.
2. A random number R, uniformly distributed between 0 and 1, is assigned to the probability factor $B(\zeta_{q_0}^2, \zeta_{br})$. The factor $B(\zeta_{q_0}^2, \zeta_{br})$ is related to the probability that q_0 will evolve from its initial scale ζ_{q_0} to a lower scale ζ_{br} (without the emission of resolvable radiation), at which point it will branch; see the discussion before equation (101) above.
3. The numerical values for $\Delta_{q_0}(\zeta_{q_0}^2, \zeta_0^2)$ and $B(\zeta_{q_0}^2, \zeta_{br})$ are substituted into equation (101), leading to the result $\Delta_{q_0}(\zeta_{br}^2, \zeta_0^2) = \Delta_{q_0}(\zeta_{q_0}^2, \zeta_0^2)/R$ and therefore to a numerical value for the second Sudakov factor $\Delta_{q_0}(\zeta_{br}^2, \zeta_0^2)$.
4. This last Sudakov factor is inverted for the value ζ_{br}, i.e. for the value of the evolution variable at which a branching occurs. If ζ_{br} is larger than the cut-off ζ_0, the branching is indeed deemed to take place; otherwise not.

If the initial quark q_0 does branch, energy fractions $z_q = z$ and $z_g = 1 - z$ are assigned to the daughter quark and gluon, respectively, according to the QCD branching probability distributions

$$[\mathcal{P}(z)]_{ij} \propto \frac{\alpha_S(p_T^2)}{2\pi}P_{i\to jk}(z), \tag{108}$$

with $P_{i\to jk}(z)$ the Altarelli-Parisi splitting kernel and where, in this case, $i = q_0$, $j = q$ and $k = g$. The naive range $0 \leq z \leq 1$ is restricted by the requirement that $E_q \geq m_{q_0}$ and $E_g \geq m_g$, with m_{q_0} and m_g as introduced above. As already discussed, the argument of α_S is the transverse momentum squared of the branching, which in terms of the evolution variable is given by $p_T \simeq z(1-z)\zeta_{br}$. When an acceptable branching has been found, initial values for the evolution variable ζ may be assigned to the two daughters, as in equation (104).

Subsequent branchings. The daughters q and g from the q_0 splitting (if it occurs) are evolved in the same manner and therefore potentially branch themselves. Thus for the gluon daughter g,

1. the Sudakov factor $\Delta_g(\zeta_g, \zeta_0)$ is calculated, where ζ_g is the initial value assigned to the evolution variable of the gluon,
2. the probability $B(\zeta_g, \zeta_{br}^g)$ that the gluon will not branch before scale ζ_{br}^g is set equal to a random number,
3. the Sudakov factor $\Delta_g(\zeta_{br}^g, \zeta_0)$ is calculated from equation (101), and then
4. inverted for the numerical value of ζ_{br}^g, which is next tested against the cut-off value ζ_0, etc.

For the evolution of this gluon, the branchings

$$g \rightarrow gg, \tag{109}$$

$$g \rightarrow q\overline{q}, \tag{110}$$

and, as a non-perturbative option (which is switched off by the default parameter settings),

$$g \rightarrow (qq)(\overline{qq}) \tag{111}$$

are considered. In these expressions, g, q and (qq) represent a gluon, a quark and a diquark, respectively.

A value for the branching scale ζ_{br}^{g} is calculated for each enabled process (109)-(111). For processes (110) and (111), a value ζ_{br}^{g} is calculated for each allowed quark or diquark flavor. The cut-off ζ_0 used in these calculations also depends on the process (109)-(111) and on the quark or diquark flavor, e.g. for process (109) ζ_0 equals twice the gluon mass while for process (110) it equals twice the quark mass, which itself depends on the flavor of the quark.

For the non-perturbative process (111) there is no Sudakov form factor. Instead, the corresponding value of ζ_{br}^{g} is chosen with a probability distribution P_{diqk}/ζ_{br}^{g} below a maximum value Q_{diqk}, where P_{diqk} and Q_{diqk} are parameters specified by the user. These parameters determine how often a gluon splits non-perturbatively into diquarks rather than into quark-antiquark (see next subsection). The default setting $Q_{diqk} = 0.0$ disables the diquark splitting option. Any diquarks generated in this process are taken to be on-mass-shell and do not undergo branching.

If the value of ζ_{br} is smaller than ζ_0 for all allowed processes, the gluon does not branch. If ζ_{br} is larger than ζ_0 for more than one process, the process having largest ζ_{br} is selected (since this branching would thereby have occurred at an earlier time in the shower development). This procedure is equivalent to having the total probability for a branching at some scale ζ (given that the gluon has not already branched at a larger scale) being equal to the sum of probabilities for each of the allowed channels. If a branching occurs, energies and evolution variables are assigned to the daughters as before, c.f. equations (108) and (104), respectively.

Each quark or gluon created in the shower is evolved in this manner until the evolution variable for each is below the appropriate cut-off ζ_0, at which point the branching stage of the HERWIG shower simulation is complete. At the end of this branching stage, each parton has been assigned an energy and a value for the evolution variable (through equations (108) and (104)). The energies are given in the boosted frame, which was introduced in order to define the initial angular variable (103) for q_0. The value of the evolution variable associated with a parton is its ***initial*** value, i.e. the value assigned to the parton at its creation (equation (104)). The 3-momenta of the partons, and in particular the correlations between those momenta, are at this stage undetermined, however.

Parton shower kinematics (second stage):
In the second stage of the HERWIG parton shower simulation, the parton 3-momenta are defined. It is at this stage that correlations between the azimuthal emission angles of the partons are introduced. It is also at this stage that final state gluons are split non-perturbatively into quark-antiquark pairs (the 'Wolfram ansatz'). Although not a

part of the parton shower proper, this ansatz will be described here so as to reflect the organization of the HERWIG computer code.

Please note that by 'final state parton' we mean any of the quarks and gluons which do not branch during the first stage of the shower simulation, i.e. before the non-perturbative gluon splittings.

Construction of 3-momenta. The construction of the parton 3-momenta begins with the calculation of virtual mass values for all partons in the shower. The mass of a ***final state*** parton is set equal to its on-shell mass value. From recursive application of the relation

$$Q_i^2 = Q_j^2 + Q_k^2 + 2z_j z_k \zeta_i^2, \tag{112}$$

(which follows from equations (102) and (103)) it is then possible to calculate the mass Q_i for each parton i (from the daughter masses Q_j and Q_k), working backwards from the final state partons to the initiating parton q_0. Note for these calculations that the values of the evolution variable assigned to j and k (i.e. $z_j\zeta_i$ and $z_k\zeta_i$) are employed.

Next, the magnitude of the 3-momenta of each parton is calculated from its energy and mass. Since the polar angles at each branch are known (through the values of the evolution variables), the 3-momenta may be resolved into their longitudinal and transverse components, relative to the parent directions. Thus it remains only to specify the directions of those transverse components for the description of the parton shower to be complete. The directions of the transverse momenta depend on the correlations between the azimuthal emission angles of the partons. In HERWIG, these correlations have two sources:

1. spin, and
2. interference due to the coherent nature of soft gluon emission.

The correlations between azimuthal emission angles due to coherence are included in the following manner. Consider the branching

$$i \rightarrow jk, \tag{113}$$

followed by

$$k \rightarrow lm, \tag{114}$$

where parton l is a gluon and where parton m may be either a quark or a gluon. The distribution of the angle between the decay planes of processes (113) and (114) is known in the soft limit, i.e. as the energy of gluon l approaches zero. The maximum of this distribution occurs when the angle between the two planes is zero. In HERWIG a simple generalization of the soft gluon limit is used to provide a smooth extrapolation to the non-soft region, see [1] for more details. To incorporate the correlation, HERWIG employs a standard Monte Carlo acceptance-rejection technique, i.e. an azimuthal angle ϕ_{test} is generated randomly with a uniform distribution between 0 and 2π. A weight is assigned by evaluating the azimuthal correlation distribution using this value ϕ_{test}. A second random number is compared to the weight to determine whether ϕ_{test} will be accepted; if not another random angle ϕ'_{test} is generated, and so on.

The correlations between azimuthal emission angles due to spin are included in an analogous fashion. For this purpose, HERWIG incorporates the results of [9,2], which provides distributions describing the angle between the decay planes of any two branchings in the parton shower due to spin, at the leading collinear logarithmic level.

For the purposes of spin correlations, parton l in (114) may be either a quark or a gluon. It is important to reiterate that spin correlations between *all* the partons in a jet are incorporated into HERWIG through implementation of the results [9,2], and that these correlations are thereby more sophisticated than a simple nearest neighbour approximation (which was used in earlier versions of HERWIG).

To simultaneously include the correlations due to both coherence and spin, a weight is determined for a randomly chosen azimuthal angle ϕ_{test} according to the distributions for coherent emission and for gluon spin separately. The total weight for this angle ϕ_{test} then equals the product of these two weights.

By working through the shower in a particular sequence, starting from the initial quark q_0 (see [9,2] for details), all azimuthal angles in the branchings are generated. Once the azimuthal angles have been determined, the full 3-momenta of the partons are calculated.

Non-perturbative gluon splitting. It is also during this second stage of the HERWIG shower simulation that final state gluons are split non-perturbatively into quark-antiquark pairs. As explained above, this splitting is performed in order to separate the two color indices of the gluon. Such a separation is necessary in order that each color index might have a uniquely associated 4-momentum, as is required by the hadronization algorithm.

The final state gluons are split as follows. Each final state gluon is assumed to decay into an up or a down quark-antiquark pair, i.e. $u\overline{u}$ or $d\overline{d}$, with a 50% probability for each. The rationale for allowing splitting only to $u\overline{u}$ or $d\overline{d}$ is that the value of the gluon virtual mass cut-off is normally too low for $s\overline{s}$ or heavier flavours to be produced at this stage. The energy of the parent gluon is divided between the quark and antiquark according to a random distribution, determined by the Altarelli-Parisi splitting kernel $P_{g\to q\overline{q}}(z)$. The polar angle of the decay is determined by the value of the energy fraction z, the azimuthal angle according to the spin correlation distribution of [9,2], the HERWIG implementation of which which was discussed in the previous subsection.

Combining showers (third stage):
A parton shower is generated, in the manner described above, separately for the jets initiated by the q_0 and by the $\overline{q}_0$ created in the electroweak scattering. The next step in the HERWIG parton shower simulation is to combine these two jets, in order that the complete e^+e^- event might be described.

Event kinematics. It is first necessary to verify that the event which has been generated is kinematically possible. Both the q_0 and $\overline{q}_0$ jets are generated with boosted energies

$$(E_{q_0})_{boosted} = \sqrt{p_{q_0}p_{\overline{q}_0}} \sim \sqrt{2}\frac{E_{CM}}{2}, \tag{115}$$

cf. eq. (107). The mass of each jet may have a value up to the energy of its initiating parton, however. Thus the mass of the combined two-jet system could be as large as $\sqrt{2}\ E_{CM}$, which exceeds the maximum permissible value E_{CM} (i.e. the mass of the original $q_0\overline{q}_0$ system).

In HERWIG, a two-jet system having a total mass larger than E_{CM} is rejected. In this case, both parton showers are regenerated, while preserving the initial conditions of the electroweak scattering.

The next to last step in the HERWIG shower simulation is to boost the 4-momenta of the partons back to the center-of-mass frame, i.e. to undo the boost introduced at the beginning of the shower evolution, so that initial angular variables (103) equal to unity for both q_0 and $\overline{q}_0$ could be defined. Note that, since the two jets have acquired effective masses by the shower evolution, the boost valid for massless jets ($\beta = 1/3$) would not give either the correct energy (E_{CM}) or vanishing longitudinal momentum for the $q_0\overline{q}_0$ jet system as a whole. Instead, the q_0 and $\overline{q}_0$ boost values are chosen in such a way that both net energy and net longitudinal momentum comes out right.

Finally, HERWIG invokes the mechanics to include an azimuthal correlation (due to the coherent emission of soft gluons) between the radiation patterns of the q_0 and $\overline{q}_0$ initiated jets (the 'cone vectors', which specify the orientations of the radiation patterns, are boosted along with the jets and aligned). For e^+e^- annihilations this jet-to-jet correlation does not exist, however, because the interfering partons q_0 and $\overline{q}_0$ are back-to-back. The parton showers initiated by q_0 and $\overline{q}_0$ are therefore simply given a random azimuthal orientation one-to-the-other.

Jet–jet interference. In the $e^+e^- \to q_0\overline{q}_0g_0$ option of HERWIG, the parton showers from the three outgoing partons are generated taking into account the coherent colour structure of the final state, to leading order in the number of colours. In this approximation, soft gluons are emitted coherently by the quark and gluon, and by the gluon and antiquark, but not directly by the quark and antiquark. This means that the initial angular variable for the quark jet is determined by the quark and gluon 4-momenta, $\xi_q = \xi_{q_0g_0}$ given by equation (103), that for the antiquark jet is $\xi_{\overline{q}} = \xi_{\overline{q}_0g_0}$, while for the gluon jet ξ_g can equal either ξ_q or $\xi_{\overline{q}}$, with equal probability.

In contrast to the $q_0\overline{q}_0$ case, the relative azimuthal orientation of the jets is now crucial, to take into account the correlations due to soft gluon coherence. This is incorporated through the mechanism of the HERWIG cone vectors. The cone vector of a jet specifies the direction of the maximum of the distribution describing the azimuthal correlation in the frame in which the parton shower is generated. This maximum lies in the plane defined by the 3-momenta of the interfering partons – say q_0 and g_0 in this case. The cone vector is boosted along with the parton shower when this latter is returned to the original laboratory frame. By rotating the q_0 and g_0 jets, as seen in the laboratory, around the q_0 and g_0 directions such that the cone vectors lie in the same plane and in the region between the q_0 and g_0, the azimuthal correlation between the two radiation patterns is included.

In the 3-jet option the angular region available to each jet is smaller than in the two-jet case, due to jet-jet interference. On the other hand, the branching probability in the event plane is enhanced by the correlations discussed above, and the gluon jet has a higher overall branching probability than a quark jet. These features lead to a satisfactory description of the characteristic properties of 3-jet events, such as the 'string effect' [10].

Cluster formation and decay:
In the leading logarithm approximations upon which the HERWIG parton shower is based, each final state quark (or gluon) usually finds itself nearby in momentum and coordinate space to a parton (or partons) having opposite color charge (or charges). Following the non-perturbative split-up of gluons, all final state partons are either quarks

or antiquarks (unless perturbative diquark production is allowed, c.f. reaction (111) above). Thus each final state parton possesses a single color index. The definition of 'nearby' in this context is formally 'of the order of the infrared cut-off scale ζ_0.' In momentum space, if nearness is measured by the combined invariant mass of a quark and its associated color-opposite antiquark, a 'near' value corresponds roughly to the rest mass values of the quarks and of the gluon, since these serve as the cut-off parameters for the shower. 'Nearby' therefore means that a quark and its associated antiquark have an invariant pair-mass of one or two GeV. Correspondingly, in coordinate space the typical invariant separation is a fraction of a fermi. The property of perturbative QCD by which final state partons may be associated pairwise into colorless, low mass *and thus hadron-like* objects is known as preconfinement [5]. These hadron-like objects are what are called clusters. Clusters are characterized by their 4-momenta and by the flavors of the quark and antiquark of which they are comprised (and by nothing else). Note that if the perturbative production of diquark pairs is allowed, a cluster may contain a diquark as one or both of its constitutive partons (although by default this possibility is not allowed).

In the following we first discuss the mechanism for cluster formation in HERWIG. The main complication here concerns clusters of large mass. Following this we describe the decay of those clusters, i.e. the HERWIG prescription for introducing hadrons.

Cluster formation. In the Monte Carlo implementation, the color flow within a parton shower is known. Therefore the identification of clusters in HERWIG is straightforward. The 4-momentum of a cluster equals the sum of the 4-momenta of its constituent quark and antiquark; note the primordial importance of the non-perturbative gluon splitting ansatz to the definition of the clusters and of their masses.

Splitting of large mass clusters. Occasionally a HERWIG cluster has a very large mass value, i.e. a quark and its associated color-opposite antiquark are not nearby in the sense defined in the introduction above. In principle, the cluster mass spectrum extends all the way up to the invariant mass of the original $q_0\overline{q}_0$ pair, i.e. up to E_{CM}. This can occur if neither q_0 nor $\overline{q}_0$ branch at all in the shower evolution. The high mass tail is therefore essentially controlled by the Sudakov form factor, and falls off more rapidly than any negative power of the cluster mass. Thus, in reality, this tail is overwhelmed by higher-twist (i.e. power-suppressed) processes that are not yet taken into account in any Monte Carlo program. However, for practical purposes, HERWIG needs to adopt a procedure for dealing with those events containing high-mass clusters. Such cases arise most frequently when the flavor of q_0 is bottom (or a heavier flavor), for which evolution through branching in the parton shower phase is less probable.

In such instances, the usual isotropic cluster decay scheme adopted by HERWIG (to be described below) is not expected to be applicable. A special anisotropic procedure is therefore invoked to split up anomalously heavy clusters, as follows.

If M_0 is the mass of cluster C_0, while m_{q_1} and $m_{\overline{q}_2}$ are the rest masses of its constituent quark and antiquark (of flavors q_1 and $\overline{q}_2$), then if the relation

$$M_0^2 \leq M_{max}^2 + \left(m_{q_1} + m_{\overline{q}_2}\right)^2 \tag{116}$$

is not satisfied, the cluster is considered to be anomalously heavy and is split. The variable M_{max} is effectively a cut on the maximum kinetic energy available to hadrons in isotropic cluster decay, and is a phenomenological parameter of the HERWIG cluster

decay model. The default value of M_{max} (denoted CLMAX in the HERWIG computer code) is 5 GeV.

With the default parameter values of HERWIG, about two per cent of all clusters do not satisfy the criterion (116) and are subjected to the cluster splitting procedure described below. As a consequence of this procedure, all high mass clusters are evolved into low mass clusters, to which the HERWIG scheme for isotropic decay to hadrons may be applied.

A cluster C_0 composed of quarks $(q_1,\overline{q}_2)$ and not satisfying relation (116) is split into two clusters of smaller mass, which we denote C_1 and C_2, composed of quarks $(q_1,\overline{q}_3)$ and $(q_3,\overline{q}_2)$, respectively. In other words, the splitting is performed by introducing a new quark pair $(q_3\overline{q}_3)$ between q_1 and $\overline{q}_2$. The flavor of this pair $q_3\overline{q}_3$ is selected to be up, down or strange, with equal probability. The rationale is that the suppression of strange quark production is purely according to phase-space for lighter clusters, and so should not be so important for the more massive ones. Mass values M_1 and M_2 for the daughter clusters C_1 and C_2 are generated through the relations

$$\begin{aligned} M_1 &= M_0 \cdot R_1^{\beta}, \\ M_2 &= M_0 \cdot R_2^{\beta}, \end{aligned} \tag{117}$$

where R_1 and R_2 are two independent random numbers, uniformly distributed between 0 and 1, and β is a phenomenological parameter (default value unity) for the daughter mass spectrum. The daughter masses thus selected are required to satisfy

$$\begin{aligned} M_0 &> M_1 + M_2, \\ M_1 &> m_{q_1} + m_{\overline{q}_3}, \\ M_2 &> m_{q_3} + m_{\overline{q}_2}, \end{aligned} \tag{118}$$

else they are regenerated.

Once the masses M_1 and M_2 of the daughter clusters are known, the magnitudes of the 3-momenta of C_1 and C_2 (and of their constituents q_1, $\overline{q}_2$, q_3 and $\overline{q}_3$) may be calculated using two-body decay kinematics. In HERWIG, the directions of all these momenta are set parallel to the direction between q_1 and $\overline{q}_2$, in the rest frame of the parent cluster C_0. Thus the cluster splitting algorithm of HERWIG has close analogies to string decay as implemented in models such as JETSET, in which a $q\overline{q}$ pair is created in the color field between a quark and its color-opposite antiquark, with but little transverse momentum relative to color field axis.

Cluster decay to hadrons. There are two principal mechanisms by which a cluster may decay to hadrons: a one-body mechanism and a two-body mechanism. The mechanism which is chosen for the decay of a particular cluster depends upon its flavor content and mass value. Suppose a cluster C is composed of a quark q_A and an antiquark $\overline{q}_B$. If the mass M_C of the cluster satisfies either relation

$$\begin{aligned} M_C &> [m(q_A,\overline{u})]_{lightest} + [m(u,\overline{q}_B)]_{lightest}, \\ M_C &> \left[m(q_A,\overline{d})\right]_{lightest} + [m(d,\overline{q}_B)]_{lightest} \end{aligned} \tag{119}$$

(where u and d are the up and the down quarks), the cluster will decay according to the two-body mechanism, which is described below.

The variable $[m(q_1,\overline{q}_2)]_{lightest}$ in equation (119) is the mass of the lightest hadron in HERWIG which has quark composition $(q_1,\overline{q}_2)$. Thus $[m(q_1,\overline{q}_2)]_{lightest}$ equals 0.1396 GeV (the mass of the π^+), if q_1 should be a u quark and $\overline{q}_2$ an anti-d quark. According to the rule (119), a cluster with flavor content $(c,\overline{s})$ must have a mass larger than $M_{D^0}+M_{K^+}=2.36\ GeV$ or else $M_{D^+}+M_{K^0}=2.37\ GeV$ in order to experience two-body decay.

A second possibility is that the cluster $(q_A,\overline{q}_B)$ does not satisfy relation (119). In this case HERWIG invokes the one-body decay mechanism, i.e. the cluster is transformed into a single particle. This single particle is chosen to be the lightest hadron having the correct flavor structure, i.e. it equals the hadron of mass $[m(q_A,\overline{q}_B)]_{lightest}$ in the notation of equation (119). A cluster with flavor $(c\overline{s})$ and a mass less than 2.36 GeV will thus be transformed into a D_s^+. The 3-momentum of this single daughter equals that of its parent; the difference in energy between cluster and hadron is transferred to a randomly selected neighbouring cluster (which, in a similar manner, has its energy but not its 3-momentum modified). If the mass of the decaying cluster is *less* than that of its daughter hadron, energy is transferred *to* the daughter from one of the neighbouring clusters. In the HERWIG implementation, these single-hadron decays are performed before the two-body decays because of the potential changes to the values of the cluster masses, upon which the two-body decays depend. Less than one per cent of all clusters decay according to this single particle scheme, however.

The two-body decay mechanism, by which most clusters decay, is the following. A quark or diquark flavor 'q_C' is selected randomly, and with equal probability, from the possibilities $d,u,s,c,b,t,uu,ud,us,ds,dd,ss$. A cluster C of flavor composition $(q_A,\overline{q}_B)$ and of mass M_C is presumed to decay into two hadrons h_1 and h_2 of flavors $(q_A,\overline{q}_C)$ and $(q_C,\overline{q}_B)$:

$$C(q_A,\overline{q}_B)\rightarrow h_1(q_A,\overline{q}_C)+h_2(q_C,\overline{q}_B). \tag{120}$$

The flavor q_C is required to satisfy the threshold cut

$$M_C > [m(q_A,\overline{q}_C)]_{lightest}+[m(q_C,\overline{q}_B)]_{lightest}\,. \tag{121}$$

For the default parameter value set, this eliminates top hadron production in the decay of non-top clusters, because of the limit on M_C; see relation (116). Similarly, if 'quark' q_A or $\overline{q}_B$ is in reality a diquark or antidiquark created in the perturbative shower (c.f. reaction (111)), the flavor of q_C is limited to the possibilities u, d, and s (to prevent the production of exotics).

Particle content of HERWIG. The identities of the decay daughters h_1 and h_2 in (120) are determined by a random selection from those hadrons, contained by HERWIG, which have the correct quark content. The hadrons contained by HERWIG are listed in Tables 5, 6 and 7. At present only one top or bottom hadron of each flavour is included, i.e. there is no Σ_b^0, no vector B mesons, etc. In addition to the hadrons listed here, HERWIG contains the lepton families (e,ν_e), (μ,ν_μ) and (τ,ν_τ). However, only hadrons are produced in cluster decay.

Phase-space factors. Hadrons h_1 and h_2, with the correct quark contents $(q_A,\overline{q}_C)$ and $(q_C,\overline{q}_B)$ for the decay (120), are selected from the hadrons listed in Tables 5, 6 and 7, with a probability proportional to their spin degeneracy. Thus an a_2^+ is five times more likely to be selected than is a π^+, should the required flavor composition be $(u,\overline{d})$.

Table 5: Meson nonets (u, d, s quarks).

0^{-+}	π^+	π^-	π^0	K^+	K^-	K^0	$\overline{K}^0$	η	η'
1^{--}	ρ^+	ρ^-	ρ^0	K^{*+}	K^{*-}	K^{*0}	$\overline{K}^{*0}$	ϕ	ω
1^{++}	a_1^+	a_1^-	a_1^0	K_1^{*+}	K_1^{*-}	K_1^{*0}	$\overline{K}_1^{*0}$	f_1	f_1'
2^{++}	a_2^+	a_2^-	a_2^0	K_2^{*+}	K_2^{*-}	K_2^{*0}	$\overline{K}_2^{*0}$	f_2	f_2'

Table 6: Heavy quark mesons.

0^{-+} charm	D^+	D^-	D^0	$\overline{D}^0$	D_S^+	D_S^-	η_c	
1^{--} charm	D^{*+}	D^{*-}	D^{*0}	$\overline{D}^{*0}$	D_S^{*+}	D_S^{*-}	J/ψ	ψ'
1^{++} charm	D_1^{*+}	D_1^{*-}	D_1^{*0}	$\overline{D}_1^{*0}$	D_S^{*+}	D_S^{*-}	χ_1	
2^{++} charm	D_2^{*+}	D_2^{*-}	D_2^{*0}	$\overline{D}_2^{*0}$	$D_{S_2}^{*+}$	$D_{S_2}^{*-}$		
0^{-+} bottom	B^+	B^-	B^0	$\overline{B}^0$	B_S^0	$\overline{B}_S^0$	B_C^+	B_C^-
1^{--} bottom	$\Upsilon(1S)$							
0^{-+} top	T^+	T^-	T^0	$\overline{T}^0$	T_S^+	T_S^-	T_C^0	$\overline{T}_C^0$
	T_B^+	T_B^-						
1^{--} top	$\Upsilon_T(1S)$							

Table 7: Baryons (including the antiparticles of the particles listed).

s=1/2 octet	p	n	Λ	Σ^+	Σ^-	Σ^0	Ξ^-	Ξ^0		
s=3/2 decuplet	Δ^{++}	Δ^+	Δ^-	Δ^0	Σ^{*+}	Σ^{*-}	Σ^{*0}	Ξ^{*0}	Ξ^{*-}	Ω^-
s=1/2 single charm	Σ_C^{++}	Σ_C^+	Σ_C^0	Λ_C^+	Ξ_C^+	Ξ_C^0	Ω_c^0	$\Xi_C'^+$	$\Xi_C'^0$	
s=3/2 single charm	Σ_C^{*++}	Σ_C^{*+}	Σ_C^{*0}	Ξ_C^{*+}	Ξ_C^{*0}	Ω_C^{*0}				
s=1/2 single bottom	Σ_b^+	Σ_b^-	Λ_b^0	Ξ_b^0	Ξ_b^-	Ω_b^-				
s=1/2 single top	Σ_t^{++}	Σ_t^0	Λ_t^+	Ξ_t^+	Ξ_t^0	Ω_t^0				

The sum of the masses of h_1 and h_2 must be less than the mass of the decaying cluster, else the combination is rejected. If this restriction on the sum of daughter masses is met, the common 3-momentum p_{CM} of the two daughters in the rest frame of the cluster parent is evaluated. p_{CM} gives the kinematical phase-space weight, based on which the cluster decay $C\rightarrow h_1+h_2$ is either accepted or rejected.

In total, the probability to select a given final state is therefore proportional to a phase-space weight W_{decay}

$$W_{decay}=p_{CM}\cdot(2S_1+1)\cdot(2S_2+1), \tag{122}$$

where S_1 and S_2 are the spin values of the two daughters h_1 and h_2. In case of rejection, the process is begun anew with the selection of a quark or diquark flavor q_C (equation (120)). If the decay channel is accepted, the daughters are given a random, uniformly distributed decay angle in the parent cluster's rest frame.

The simple decay algorithm outlined above ensures that additive quantum numbers are conserved in cluster decay, while baryon and heavy flavour production are naturally suppressed in proportion to the phase-space factor p_{CM}. The algorithm does not guarantee isospin conservation, however. For example, the range of hadrons available with flavour composition $u\overline{u}$ is different from that for $u\overline{d}$, so there is no mechanism to ensure that precisely equal numbers of π^0 and π^+ mesons are produced. In practice, there is equality at about the 10% level, with somewhat worse agreement in rare cases (e.g. around 30% more Δ^{++} than Δ^+ baryons). This could be corrected, at considerable cost in computer time, by comparing directly all possible two-hadron decays of a given cluster, without going though the simplifying intermediate step of selecting a produced flavour q_C.

Decays:
Leptons and hadrons are produced in HERWIG either through the decay of clusters or through the decay of hadrons. The manner in which a hadron decays depends on the flavor of its heaviest quark. The decays of light hadrons (i.e. hadrons only containing u, d and s quarks) and charm hadrons are treated in a straightforward manner using tables. The decays of hadrons containing bottom or top quarks are treated through recourse to the HERWIG parton shower and cluster decay formalism, thus permitting the possibility of gluon radiation. The decay of the τ lepton is also treated using this parton shower and cluster decay technique.

Light and charm hadron decays. To specify which light and charm hadrons are unstable and how they will decay, HERWIG employs lookup tables. A first table gives the number of decay modes for each charm or light hadron. A value of zero in this table means that a particle is stable. Unstable particles have entries in a second table, which states the branching ratio for each of its channels. In the Monte Carlo simulation, a decay channel is selected randomly with a weight proportional to its branching ratio. A third table contains the identities of the decay daughters. Note that in HERWIG a maximum of three decay daughters is allowed. Thus many particles decay according to a simplified scheme, which is not always in agreement with experimental data. For example, important four- and five-body non-resonant decays have been measured for the D^0 and D^+; these channels are not explicitly included by HERWIG. In some cases the HERWIG decay tables also exclude measured decay channels even should the number of daughters be three or less, e.g. the π^0 Dalitz decay, $\pi^0 \to e^+e^-\gamma$ (B.R. = 1.2%), is missing. Instead the π^0 decays 100% of the time to $\gamma\gamma$ in the HERWIG simulation.

The kinematics of these decays is determined by either two- or three-body Lorentz invariant phase-space, as dictated by the number of daughters. An exception occurs for the neutron and for the muon, for which the daughter 4-momenta are subjected to an acceptance-rejection procedure in order that they be weighted according to a V-A matrix element. This is the only decay matrix element incorporated into HERWIG, however (the muon and neutron are in any case stable in the HERWIG algorithm, so this point isn't particularly relevant). Similarly there is no mechanism which incorporates momentum correlations in hadron decays due to spin.

Table 8: Branching ratios for weak decays of bottom and top hadrons and of the τ.

	$u\overline{d}$ or $d\overline{u}$	$c\overline{s}$ or $s\overline{c}$	$e^-\overline{\nu}_e$ or $e^+\nu_e$	$\mu^-\overline{\nu}_\mu$ or $\mu^+\nu_\mu$	$\tau^-\overline{\nu}_\tau$ or $\tau^+\nu_\tau$
b quark	0.55	0.20	0.11	0.11	0.03
t quark	0.34	0.33	0.11	0.11	0.11
τ lepton	0.64	0.0	0.18	0.18	0.0

Heavy hadron and τ lepton decays. The decays of bottom and top hadrons, and of the τ lepton, are treated differently from the decays of light and charm hadrons. In particular, the HERWIG parton shower algorithm is utilized, to provide a possibility for quarks to radiate.

To describe the decay of a heavy quark hadron H, HERWIG first resolves that hadron into its quark constituents Q and $\overline{q}_s$, where Q is the heavy quark flavor (bottom or top) and $\overline{q}_s$ is the spectator, either $\overline{u}$, $\overline{d}$, $\overline{s}$, $\overline{c}$, uu, ud, us, dd, ds or ss. A 4-momentum is assigned to Q and $\overline{q}_s$ by the relations

$$p_Q = \frac{m_Q}{m_H} \cdot p_H,$$
$$p_{\overline{q}_s} = \frac{m_{\overline{q}_s}}{m_H} \cdot p_H, \tag{123}$$

with m_Q and $m_{\overline{q}_s}$ the rest masses of the quarks, and with m_H and p_H the mass and 4-momentum of the heavy hadron H. Note that with these assignments, Q and $\overline{q}_s$ are placed approximately on their mass shells.

Next, the heavy quark Q is presumed to decay weakly, i.e.

$$Q \to q_d + W \to q_d + (f_1\overline{f}_2), \tag{124}$$

while $\overline{q}_s$ merely acts as a spectator. In the weak decay (124), Q and q_d are assumed sequential, and there is no mixing between quark families in the $f_1\overline{f}_2$ system. Thus if Q is a bottom quark, its daughter quark q_d is always a charm quark. The identity of the fermion pair $(f_1\overline{f}_2)$ is selected randomly from a table of branching ratios, see Table 8. The 4-momenta of the daughters q_d, f_1 and $\overline{f}_2$ in (124) are given by phase-space weighted by a V-A matrix element. An additional weight on the invariant mass of f_1 and $\overline{f}_2$ (i.e. on the invariant mass of the W boson in (124)) incorporates the effects of the W propagator, whether real or virtual.

The HERWIG parton shower formalism is invoked to describe the development of the two color-singlet systems, $(f_1\overline{f}_2)$ and $(q_d\overline{q}_s)$. There is no mixing allowed between these two systems, since they are not colour-connected. We next discuss the evolution of the $(f_1\overline{f}_2)$ system, assuming that f_1 and $\overline{f}_2$ are partons. Following this we describe the case should they be leptons, and the evolution of the spectator system $(q\overline{q}_s)$.

The $(f_1\overline{f}_2)$ system is treated in an identical manner to the initial quark-antiquark $(q_0\overline{q}_0)$ pair of the e^+e^- event. Thus f_1 and $\overline{f}_2$ each initiate a separate jet of partons. To generate these jets, initial values for the angular and evolution variables are assigned to f_1 and $\overline{f}_2$ through the relations (103) and (102). The parent i in the first branching

of these showers is the W boson of (124), which thereby takes the place that the γ/Z^0 boson has for the $q_0\overline{q}_0$ system. As usual, the evolution of f_1 and of $\overline{f}_2$ is normally performed in the laboratory frame of the original $e^+e^- \to q_0\overline{q}_0$ event. In contrast to the case of the $q_0\overline{q}_0$ system, the initiating partons f_1 and $\overline{f}_2$ are not (ordinarily) back-to-back. Therefore, the radiation patterns from the two showers have to be oriented one-to-the-other, in order to simulate the azimuthal correlations between the two jets due to the coherent emission of soft gluons (this in addition to the azimuthal correlations between the partons *within* each jet due to coherence and spin).

The azimuthal correlation between the f_1 and $\overline{f}_2$ jets is incorporated through the mechanism of the HERWIG 'cone vectors', as explained previously for 3-jet final states. By rotating the f_1 and $\overline{f}_2$ jets, as seen in the laboratory, around the f_1 and $\overline{f}_2$ directions, such that the cone vectors lie in the same plane and in the region between the f_1 and $\overline{f}_2$, the azimuthal correlation between the two radiation patterns is included.

Once the parton showers from f_1 and $\overline{f}_2$ have been combined, the HERWIG algorithms for cluster formation and decay are applied in order to describe the transition of the $(f_1\overline{f}_2)$ system to hadrons.

Should the fermions f_1 and $\overline{f}_2$ from the heavy quark decay be leptons, i.e. $(f_1\overline{f}_2)=(e\nu_e)$, $(\mu\nu_\mu)$ or $(\tau\nu_\tau)$, they do not of course initiate parton showers. Instead they are merely transported through the parton shower machinery but without inducing radiation. They thus appear in the HERWIG event record with their momenta as given by the V-A decay simulation, eq. (124).

The daughter-spectator system $(q_d\overline{q}_s)$ from the heavy hadron decay is evolved in a similar manner to the $(f_1\overline{f}_2)$ system, but with the following difference: since all the parton branching in this system comes from the disappearance of the heavy quark Q and the creation of its daughter q_d, the spectator $\overline{q}_s$ does not radiate and thus remains on-shell. As in equations (105) and (106), the initial evolution variables of the Qq_d system satisfy

$$\zeta_Q\zeta_{q_d} = p_Q p_{q_d} \tag{125}$$

so the choice $\zeta_Q = m_Q$, $\zeta_{q_d} = p_Q p_{q_d}/m_Q$ is made, which suppresses all 'pre-radiation' from the heavy quark and assigns all emission to the outgoing q_d jet. See [11] for a full discussion of the treatment of coherent radiation in heavy flavour processes in HERWIG.

The HERWIG simulation of τ lepton decays is less complicated than its simulation of heavy quark hadron decays, because there is no spectator system to evolve. Thus for the treatment of the τ, the weak decay

$$\tau \to \nu_\tau + W \to \nu_\tau + (f_1\overline{f}_2) \tag{126}$$

takes the place of the decay (124). The identity of the pair $(f_1\overline{f}_2)$ is selected randomly, according to the probabilities in Table 8. Should $(f_1\overline{f}_2)$ be a quark pair, $(f_1\overline{f}_2)=(d\overline{u})$, the transition to hadrons is performed by applying the HERWIG parton shower and cluster algorithms in the manner described above for the decay of the $(f_1\overline{f}_2)$ system in heavy quark decays.

4.7.4 Discussion of Certain Model Elements

The HERWIG Monte Carlo offers an elaborate simulation of parton showers and a simple model of hadronization phenomena. It is unequaled by any other parton shower Monte

Carlo, at the time of this writing, in the completeness of its simulation for azimuthal correlations due to the coherent emission of soft gluons and due to spin, for example.

Some elements of HERWIG are radically different from other programs and may seem strange on that account. We discuss some of these elements further in this section.

Perhaps the most distinctive feature of the hadronization model used in HERWIG is the non-perturbative gluon splitting procedure, the so-called 'Wolfram ansatz' [4]. This gluon splitting is carried out so that each color index might have an individually defined 4-momentum, which itself is necessary for the definition of the cluster masses. As a consequence of the Wolfram mechanism, each gluon which does not split into a quark-antiquark pair during the perturbative phase does so afterwards. In turn this means that each gluon gives rise to the creation of one new cluster and thus of at least one hadron (and usually of at least two hadrons).

It is sometimes suggested that gluon splitting implies an 'infrared instability', since the event configuration is sensitive to the addition of a soft gluon in the final state. For example, in $e^+e^- \to q_0\overline{q}_0$ with no gluon emission, the final state consists of a single static cluster of high mass, whereas the emission of one soft gluon could convert this to two fast low-mass clusters.

There is in fact no theorem in QCD which states that the configuration of a particular *exclusive* final state should be stable under addition of soft gluons. Indeed one expects very large fluctuations in the colour structure of such states. The factorization theorems that imply such stability refer to *inclusive* cross-sections, in which one sums over large classes of final states. The Wolfram ansatz for hadronization does not violate any of these theorems, or any well-founded intuition about QCD. On the other hand, hadronization mechanisms like the string model, which impose stability under soft gluon emission at the exclusive level, would seem to suppress the long-range colour fluctuations that would be expected to arise from soft gluon effects.

A second element of HERWIG which draws attention concerns the flavor selection for non-perturbative quark pairs. Quark pairs are produced non-perturbatively at three stages in the HERWIG hadronization scheme:

Model Stage	flavor of $(q_f\overline{q}_f)$ pair
(I) Non-perturbative gluon splitting $g \to q_f\overline{q}_f$	$(u\overline{u})$ or $(d\overline{d})$
(II) Heavy cluster splitting $C(q_1,\overline{q}_2) \to C(q_1,\overline{q}_f) + C(q_f,\overline{q}_2)$	$(u\overline{u})$, $(d\overline{d})$ or $(s\overline{s})$
(III) Cluster decay (to hadrons) $C(q_1,\overline{q}_2) \to h(q_1,\overline{q}_f) + h(q_f,\overline{q}_2)$	$(u\overline{u})$, $(d\overline{d})$, $(s\overline{s})$, $(c\overline{c})$, $(b\overline{b})$, $(t\overline{t})$ or a diquark pair

As remarked earlier, the limitation at stage (I) is explained by the authors to be a consequence of the lack of phase-space for heavier flavours, for parameter values in the normal range. The cluster splitting (II) could be extended to generate heavier flavours (in proportion to phase-space), but since it is applied only to a relatively small fraction of clusters this would most likely make little difference. Thus the flavour composition of the final state is in practice controlled by the perturbative phase (primary hard process and parton branching) and by hadronization stage (III).

A third element which we should discuss further concerns the Lorentz invariance of the parton shower algorithm. In HERWIG, the individual parton showers of e^+e^- annihilations are not generated in the event center-of-mass but in boosted frames in which the quark and antiquark energies are scaled up by a factor of approximately $\sqrt{2}$. The rescaling of the jet energy is compensated by a corresponding reduction of the angular variable $\sqrt{\xi}$ of the jet, and is performed so that the starting value of this quantity is not greater than unity. Since the parton branching depends only on the combination $E\sqrt{\xi}$, the structure of the parton shower (i.e. the branching topology and energy fractions at each branching) is insensitive to this boost. However, when the full kinematics are reconstructed, some frame dependence is inevitably introduced. This is in fact a problem common to all current approaches to jet evolution, since the energy or momentum fraction z can never be given a Lorentz-invariant definition. This is because the relevant Feyman graphs have to be evaluated in a non-covariant (axial) gauge and z is the momentum fraction along the direction of the gauge vector. (Sometimes it is claimed that the definition of z as the $E + p_L$ fraction, where p_L is the longitudinal momentum, is invariant, but this is true only for boosts in the longitudinal direction, and this definition is not rotationally invariant.) In HERWIG, z is the energy fraction in the boosted frame, which is rotationally invariant but introduces some boost dependence for slow partons. This does not significantly affect any quantities that are dominated by leading collinear and/or infrared logarithms, but it can affect quantities that are not really leading-logarithmic.

In practice there is no problem of Lorentz invariance for the bulk of events which are generated, but there may be some frame dependence in the numbers of events generated with a clear multi-jet structure. Thus the number of reconstructed 4-jet events observed in a HERWIG event sample may differ depending on the value of the boost factor which is chosen for the generation of the showers [12]. It should be emphasized, however, that this problem – if it exists – is of relatively small importance. From a theoretical standpoint it is not expected that HERWIG should describe the reconstructed 4-jet rate in a strictly consistent manner, because this rate is not determined by leading logarithms. From a practical (experimental) standpoint, an uncertainty in the reconstructed 4-jet rate is relatively unimportant, because 4-jet events provide only a small contribution to the global event distributions, and the predictions for the reconstructed 4-jet events themselves are unaffected – for it is mainly the *number* of 4-jet events which is affected, not their internal characteristics.

More serious criticisms can be made of the simplified models employed by HERWIG for cluster and hadron decays. Primarily in order to save time, the 'phase-space' decay model for a $q_i\overline{q}_j$ cluster actually proceeds through the intermediate step of randomly choosing a flavour f (a quark or diquark) and then choosing the identities of the trial decay hadrons $q_i\overline{f}$ and $f\overline{q}_j$. This leads to branching fractions that do not precisely follow phase-space, and in particular do not conserve isospin to better than about 10% on the average. In the hadron decay model, light and charm hadrons are limited to at most three decay daughters. This restriction is at variance with experiment, by which important four- and five-body decay channels have been measured, and in which there is no observation of resonance structure (as an example consider the decay $D^+ \to K^-\pi^+\pi^0e^+\nu_e$; B.R.=4.4% [13]). Such decays cannot be explicitly incorporated into HERWIG. The decay tables are not structured to handle them, nor are there decay routines to generate the necessary multi-body phase-space. Furthermore,

Table 9: Principal parameters of the HERWIG Monte Carlo.

Parameter Name	Default Value	Meaning
PART1	'PBAR'	Colliding particle type 1
PART2	'P '	Colliding particle type 2
PBEAM1	900.	Momentum of particle PART1 (GeV/c)
PBEAM2	900.	Momentum of particle PART2 (GeV/c)
IPROC	1705	Scattering process code
QCDLAM	0.20	Λ_{QCD}, the scale parameter (GeV)
NFLAV	6	Number of flavors
RMASS(1)	0.32	Down quark mass (GeV)
RMASS(2)	0.32	Up quark mass (GeV)
RMASS(3)	0.50	Strange quark mass (GeV)
RMASS(4)	1.80	Charm quark mass (GeV)
RMASS(5)	5.20	Bottom quark mass (GeV)
RMASS(6)	50.0	Top quark mass (GeV)
RMASS(13)	0.65	Gluon mass (GeV)
RMASS(200)	92.0	Z^0 mass (GeV)
CLMAX	5.0	Cluster mass cut-off parameter (GeV)
SWEIN	0.23	$\sin^2\theta_W$; θ_W= weak mixing angle
DMZ	2.80	Z^0 width (GeV)
QDIQK	0.0	Diquark production scale (GeV)
AZSOFT	.TRUE.	Correlations due to coherence
AZSPIN	.TRUE.	Correlations due to spin

HERWIG does not implement correlations between the momenta of decay products in hadron decays, due to spin or to due to decay matrix elements. Also on the issue of particle decays, it could be stated that the decay simulation for the τ is unnecessarily complicated and model dependent, given that the τ has been essentially 100% reconstructed experimentally in exclusive decay channels. This would suggest that the hadronic decays of the τ should be handled through lookup tables rather than through the parton shower formalism.

4.7.5 Main Parameters

The principal parameters of the HERWIG Monte Carlo which are of interest from the point of view of an e^+e^- annihilation experiment are given in Table 9.

The parameters PART1, PART2 and IPROC state the type of interaction to be simulated. The default values shown in Table 9 for these parameters are those which are given in the example program which the authors of HERWIG provide with their computer code. For e^+e^- annihilations with a normal mixture of flavors, these parameters take the values
PART1 = 'E+ ',
PART2 = 'E- ', and
IPROC = 100.

To generate final states from quark-antiquark pairs of flavour IQ (IQ = 1–6 for d, u, s, c, b, t in that order), one sets IPROC = 100+IQ. Corresponding 3-jet final states are generated by IPROC = 110 or 110+IQ.

The magnitudes of the 3-momenta of the colliding particles PART1 and PART2 are assigned to the variables PBEAM1 and PBEAM2.

The parameters QCDLAM, RMASS(13) and CLMAX are the main user controlled variables for the adjustment of HERWIG predictions to experimental data. The parameter QCDLAM is the QCD scale parameter Λ as implemented in HERWIG; its value therefore primarily controls the likelihood for parton branching in the shower. Note that QCDLAM cannot be identified with $\Lambda_{\overline{MS}}$ or with the value of Λ calculated using some other renormalization scheme. This is because the leading collinear and infrared logarithmic formalism on which the HERWIG parton shower is based leaves this renormalization scheme ambiguous. The parameter RMASS(13) is a formal mass value given to the gluon and serves as the main cut-off parameter for the parton shower. CLMAX is the threshold parameter whose value determines whether a cluster will evolve by string-like splitting rather than by direct decay to hadrons, c.f. relation (116).

The quark masses RMASS(1) – RMASS(6) and the diquark masses serve as the other cut-off parameters to the perturbative branchings. With the exception of the *top* quark mass RMASS(6), their values are not normally changed by the user. The masses of the diquarks are calculated internally by HERWIG using the quark masses and neglecting binding energies, e.g. the mass of the (us) diquark is assigned the value (RMASS(2)+RMASS(3)) during the HERWIG initialization phase.

The parameter QDIQK states the maximum value of the HERWIG evolution variable at which a gluon may split perturbatively into a diquark pair, c.f. relation (111). Thus, with a default value equal to *zero*, perturbative diquark production is not allowed.

Note that top quark production *is* allowed by default (NFLAV = 6), should such production be kinematically possible.

The three parameters RMASS(200), SWEIN and DMZ determine the electroweak cross-sections and asymmetry angles of the initial quark pair $q_0\overline{q}_0$, c.f. equation (98). The variable DMZ is not a general parameter of the HERWIG model but is defined locally ***in-subroutines-as-needed*** through data statements.

The logical variables AZSOFT and AZSPIN enable or disable azimuthal correlations between partons within a jet, due to the coherent emission of soft gluons or due to spin, respectively. If AZSOFT is '.TRUE.', then soft gluon azimuthal correlations ***between*** jets are included as well, where appropriate, but in the current version this is not yet implemented for spin correlations. These do not exist anyway for light quark production in e^+e^- annihilation, due to helicity conservation.

4.7.6 Comparison with Data

The adaptation of the HERWIG Monte Carlo to the simulation of e^+e^- annihilations is quite recent. We know of no published results from experimental groups which compare e^+e^- annihilation data to the predictions of HERWIG. However, the program Bigwig [14] is written by the authors of HERWIG and is one of its direct predecessors. Bigwig has been extensively compared with e^+e^- annihilation data by experimental groups [15]; some of the main results are discussed in section 5 of this report. The parton shower simulations in Bigwig and in HERWIG are quite similar, while the fragmentation models (cluster formation and decay) are essentially identical. Therefore it is reasonable to expect that the predictions of HERWIG will not much differ from those of Bigwig in their essential characteristics. The predictions of Bigwig are generally found to be in good agreement with high energy e^+e^- annihilation data, both with regard to the global multi-hadronic event shapes [15] and with regard to the production rates and inclusive energy spectra of individual hadron species [16]. This is quite an impressive achievement for a model with so few adjustable parameters.

The authors of HERWIG have compared the predictions of their model to experimental data from e^+e^- annihilations, pp collisions and $p\overline{p}$ collisions. Some aspects of HERWIG specific to these latter two processes, such as space-like parton branchings, have not been discussed here. The predictions of HERWIG were compared to e^+e^- annihilation data for the average charged multiplicity as a function of center-of-mass energy (up to 46 GeV) and for some global multi-hadronic event distributions (at 29 GeV). For pp data, HERWIG's predictions for the transverse momentum distribution of μ pairs produced through the Drell-Yan mechanism at $\sqrt{s}$=63 GeV were tested. Data from $p\overline{p}$ collisions at $\sqrt{s}$=630 GeV were compared to the predictions of HERWIG for the transverse momentum distribution of W bosons and for the transverse energy spectrum of jets. In all cases the authors conclude that their model provides an adequate description of data; see reference [1] for further details.

4.7.7 Installation and Availability

The HERWIG Monte Carlo is available in FORTRAN computer code either from the authors or from the CERN program library. To obtain a copy from the authors, BITNET mail may be addressed to BRW1@UK.AC.CAMBRIDGE.PHYSICS.HEP or DECNET mail to 19616:WEBBER from inside the UK, and BRW1@PHY-HEP.CAM.AC.UK from outside. Alternatively, those with DECNET access can copy the latest version, together with documentation and test program output, from the CERN VAX directory VXCERN::DISK$CR:[WEBBER.HERWIG]. The version from the authors is written in VAX FORTRAN but is convertable to standard FORTRAN 77 with relatively little effort (the main VAX specific elements are VAX include statements for common block insertion). An advantage of obtaining a copy from the authors is that the CERN library version is updated only occasionally.

One problem which has been noted with the installation of HERWIG is caused by use of generic names for common blocks such as 'TRACK' and 'EVENT'. This practice has led to name conflicts with the common block or subroutine names of other libraries when HERWIG is included in a large software environment, such as a detector simulation and reconstruction program. This problem has required users to rename the common blocks in HERWIG to something unique (as the simplest solution).

As a last comment we note that HERWIG requires a relatively large initialization time because tables of Sudakov factors are calculated at begin-run. These tables are used for interpolation during event generation in the basic decisions on parton branchings, c.f. equation (101). This initialization time is 3 minutes on a VAX 2000, 80 seconds on an Apollo DN-3500 but just several seconds on an IBM 3090. There is an option to write the form factors onto disk and then to re-read them for subsequent runs which use the same parameter values, in order to avoid this initialization when possible. The event generation rate for e^+e^- annihilations at 92 GeV is about 5000 events per hour on a VAX 2000 or on an APOLLO DN-3500.

References

[1] G. Marchesini, B.R. Webber, Nucl. Phys. **B310** (1988) 461

[2] I.G. Knowles, Nucl. Phys. **B310** (1988) 571

[3] For reviews, see:
A. Bassetto, M. Ciafaloni, G. Marchesini, Phys. Rep. 100 (1983) 201;
Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, A.H. Mueller, Rev. Mod. Phys. **60** (1988) 373;
B.R. Webber, Ann. Rev. Nucl. Part. Sci. **36** (1986) 253

[4] S. Wolfram, in Proc. 15th Rencontre de Moriond (1980), ed. J. Tran Thanh Van;
G.C. Fox, S. Wolfram, Nucl. Phys. **B168** (1980) 285

[5] D. Amati, G. Veneziano, Phys. Lett. **B83** (1979) 87

[6] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Z. Physik **C27** (1985) 65

[7] D. Amati, A. Bassetto, M. Ciafaloni, G. Marchesini, G. Veneziano, Nucl. Phys. **B173** (1980) 429

[8] A.H. Mueller, Phys. Lett. **B104** (1981) 161;
A. Bassetto, M. Ciafaloni, G. Marchesini, A.H. Mueller, Nucl. Phys. **B207** (1982) 189;
B.I. Ermolaev, V.S. Fadin, JETP Lett. **33** (1981) 285;
Yu.L. Dokshitzer, V.S. Fadin, V.A. Khoze, Z. Physik **C15** (1983) 325, *ibid.* **C18** (1983) 37

[9] M.P. Shatz, Nucl. Phys. **B224** (1983) 218;
B.R. Webber, Phys. Lett. **B193** (1987) 91;
I.G. Knowles, Nucl. Phys. **B304** (1988) 767

[10] JADE Collaboration, W. Bartel *et. al.*, Phys. Lett. **B134** (1984) 275, **B157** (1985) 340;
TPC Collaboration, H. Aihara *et. al.*, Z. Physik **C28** (1985) 31, Phys. Rev. Lett. **57** (1986) 945;
Mark II Collaboration, P.D. Sheldon *et. al.*, Phys. Rev. Lett. **57** (1986) 1398

[11] G. Marchesini, B.R. Webber, Cambridge preprint Cavendish-HEP-88/9 (1988)

[12] TASSO Collaboration, private communication

[13] Particle Data Group, G.P. Yost *et. al.*, Phys. Lett. **B204** (1988) 1

[14] G. Marchesini, B.R. Webber, Nucl. Phys. **B238** (1984) 1, *ibid.* **B238** (1984) 492

[15] For recent studies, see:
Mark II Collaboration, A. Petersen *et. al.*, Phys. Rev. **D37** (1988) 1;
TASSO Collaboration, W. Braunschweig *et. al.*, Z. Physik **C41** (1988) 359

[16] For a recent review, see:
W. Hofmann, Ann. Rev. Nucl. Part. Phys. **38** (1988) 279

4.8 JETSET

4.8.1 Basic Facts

Program name: JETSET, The Lund Monte Carlo for Jet Fragmentation and e^+e^- Physics [1,2]

Version: JETSET 7.1, from January 1989; some additions in May 1989

Author: Torbjörn Sjöstrand
CERN/TH
CH-1211 Geneva 23, Switzerland
Phone (Switzerland) + 22 - 767 28 20
BITNET TORSJO @ CERNVM
One routine (LUSHOW) mainly written by Mats Bengtsson.
Theory input from B. Andersson, M. Bengtsson, G. Gustafson, T. Sjöstrand, and B. Söderberg.

Program size: 9278 lines

4.8.2 Physics Introduction

JETSET has its origin in the efforts of the Lund theory group to understand the fragmentation process. The basic idea is that the confinement potential of QCD corresponds to a colour flux tube which is stretched between the partons of a colour singlet state, and that the time evolution of this flux tube can be represented by the dynamics of the massless relativistic string. The presumed (effective) one-dimensional nature of the QCD confinement force is thus included in the model by construction. In this picture, a quark or antiquark corresponds to an end of the string, while gluons correspond to kinks on the string. Thus a gluon is attached to two string pieces, reflecting its double colour charge. Both the string itself and the endpoints (quarks) and kinks (gluons) carry energy and momentum.

The string may break by the production of $q'\overline{q}'$ pairs, such that an initial $q\overline{q}$ colour singlet is split into two colour singlets, $q\overline{q}'$ and $q'\overline{q}$. It is presumed that the $q'\overline{q}'$ pair production is described by semiclassical tunnelling from the vacuum. The original string is, by this process, broken up into pieces of smaller and smaller mass. The sharing of energy and momentum at the breakups is determined from the concept of left-right symmetry, which states that the overall picture should look the same irrespectively of the order in which causally disconnected breakups are considered. The breaking is assumed to stop when the masses of the string segments reach the hadronic mass scale. The resulting $q_i\overline{q}_j$ string states are associated with mesons of the corresponding flavour content. Baryon production is introduced by allowing the production of diquark-antidiquark pairs, in addition to quark-antiquark pairs, to describe the string breakup. The requirement that hadrons be produced on mass-shell imposes correlations between the space-time production points of the different quark-antiquark or diquark-antidiquark pairs; this feature is used to implement an iterative structure for the fragmentation.

The Lund string fragmentation concept was developed in the years around 1980 [3]. Evolution of model and program went hand in hand. The finished form of the model was reached around 1983 [4,5,6]; since then there has been no significant modification to the string fragmentation concept. The development of the program has continued unabatedly, however. In the process of exploring the differences between string and

other fragmentation concepts [7], several independent fragmentation algorithms were introduced into JETSET so as to provide alternatives to the string fragmentation default. Over the years, the program has also evolved to give a fairly detailed description of particle decays and to contain many utility routines. The production of a primary parton configuration to simulate e^+e^- annihilations was incorporated in order to facilitate the comparison of the string model with data. Use was made of matrix elements, eventually to second order, and with γ/Z^0 production and initial state QED radiation included, an evolution again completed around 1983.

The main development in recent years has been the addition of parton shower evolution [8,2], which has replaced the matrix element approach as the default option. The JETSET algorithm is of the coherent branching type, for which parton branchings are required to take place with decreasing emission angles as the shower evolves. Azimuthal anisotropies due to gluon polarization are also included. In its details, the algorithm is (conceptually) less elegant than some of the others covered in this report. On the other hand, a serious attempt has been made in JETSET to describe hard gluon emission well. This is important since hard gluons play a major role in determining the global features of events.

The JETSET program version described here is 7.1, which is of fairly recent date. Differences compared to its immediate predecessor, version 6.3, are small from a physics point of view, but there is no backwards compatibility, in particular because of a change from the old JETSET particle numbering scheme to the new PDG codes (see section 6.1) and because of changes to the common block structures. It should be noted that the expression 'Lund model' only refers to the basic string fragmentation concept, and not to any actual program. Thus, while one may speak of 'Lund fragmentation', it is incorrect to refer to 'Lund 7.1' (or, worse still, 'LUND 7.1'; after all, Lund is a place, not an abbreviation).

The main user-called routine for e^+e^- event generation (LUEEVT) acts as a driver. The flow inside this program corresponds rather well to the subheadings of the detailed physics description, section 4.8.3, and therefore it will be outlined here.

1. In connection with the first event of a run, some initialization is normally performed for the total cross-section and, if initial state radiation is switched on, for the γ spectrum.
2. For the initial state radiation option, a γ may be generated. The partonic (i.e hadronic) CM energy, after any γ emission, is calculated.
3. The flavour of the primary $q\overline{q}$ pair is selected (at the partonic CM energy).
4. In the matrix element option, a selection is made between the generation of a two-, three-, or four-jet event. For parton showers, only two jets are generated.
5. The partons are oriented in angle, still in the partonic CM frame.
6. If an initial state γ was emitted, the partonic state is boosted and rotated to the overall CM frame.
7. In the parton shower option, a shower is evolved from the primary $q\overline{q}$ pair.
8. Partons are arranged in the order in which they appear along the string.
9. If a colour singlet system has a very small mass at this stage, an attempt is made to form two hadrons. If that fails, one hadron is formed, and energy and momentum are shuffled elsewhere in the event.

10. The partonic system is fragmented using string or, as an option, independent fragmentation.
11. The decay chains of unstable particles are generated. After the decay of short-lived resonances, but before that of long-lived particles, it is possible to optionally include Bose-Einstein effects according to a simple algorithm.
12. Additional partons may be produced through the decays of heavy particles such as top hadrons. For these "secondary partons," the process starts over at point 7, with new showers, followed by more fragmentation and decays.

The modular structure outlined above makes it easy to adapt other Monte Carlos to the framework of the JETSET library. Of the programs covered in this report, ARIADNE, NLLjet and TIPTOP use the JETSET fragmentation/decay package; ARIADNE in addition makes use of several of the JETSET routines specific to e^+e^- annihilations. During the course of this workshop, an interface was written to the Higgs event generator ($e^+e^- \to (\gamma)Z^0H^0$) of Kleiss [9]. Many similar interfaces have been written inside experimental collaborations. Other programs based on JETSET include PYTHIA for hadron collisions [10] and LEPTO for leptoproduction [11]. JETSET, ARIADNE, PYTHIA and LEPTO are all examples of 'Lund Monte Carlos', a concept considerably broader than that of 'the Lund model'.

4.8.3 Detailed Physics Description

Electroweak cross-section:
Cross-sections are given with full γ/Z^0 electroweak structure included to lowest order (with radiative corrections as described in the next section), times second order QCD final state corrections. This corresponds to the equations of section 2.1.1, except that the last term in R_{QCD} is not implemented. Quark mass effects are by default included, i.e. there is a factor $v_q(3-v_q^2)/2$ multiplying the vector current terms (electromagnetic and weak) and a factor v_q^3 the axial current ones (weak only). Here v_q is the velocity of the quark, $v_q = \sqrt{1-4m_q^2/E_{CM}^2}$. The mass dependence can optionally be replaced by a step function change when passing a heavy flavour threshold, with no mass suppression at all above threshold. For simplicity, the threshold is taken to be at $2m_q + 1$ GeV, which gives a few hundred MeV safety margin compared to twice the mass of the lightest hadron (which, in JETSET, is typically $m_q + 300$ MeV).

The standard model input parameters are α_{em}, $\sin^2\theta_W$, m_Z, Γ_Z, and $\Lambda_{\overline{MS}}$ (the latter used to calculate a second order α_S). Alternatively, the Z^0 width may be calculated at initialization, based on a three-generation scenario, with top excluded by default.

The incoming e^+ and e^- may be transversely and/or longitudinally polarized. Four input parameters may be given: the longitudinal polarization of the electron and positron separately, the (geometric mean) transverse polarization, and the (average) azimuthal angle of transverse polarization.

Initial state QED radiation:
Initial state photon radiation, by default switched off, can be included to lowest, unexponentiated order, following the approach developed by Berends, Kleiss and Jadach [12].

In particular, the photon spectrum is given by the Bonneau-Martin [13] type formula

$$\frac{d\sigma}{dk} = \frac{\alpha_{em}}{\pi}\left(\ln(s/m_e^2)-1\right)\frac{1+k^2}{k}\sigma_0(s'=(1-k)s). \tag{127}$$

Here $k = 2E_\gamma/E_{CM}$ is the fractional photon energy, $s = E_{CM}^2$, s' is the hadronic CM energy-squared after photon emission, and $\sigma_0(s')$ the lowest order cross-section at the reduced energy. The soft photon cut-off is at 1% of the beam energy, the hard photon one at 99%. Soft and virtual photon effects below the 1% cut-off are included in the no-emission class of events.

First order QED loops (vacuum polarization) are also included, except that, for the hadronic vacuum polarization contribution to the γ propagator, the BKJ lookup table [12] is replaced by a simple parametrization, $4.578 \cdot \ln(s/0.932\ \mathrm{GeV}^2)$, in units where e.g. the muon contribution is $\ln(s/m_\mu^2) - 1$. There are no weak loop effects included in the program.

In addition to the photon energy, it is necessary to select a polar and an azimuthal angle for the emitted photon, as well as an additional angle characterizing the recoiling hadronic system [12]. This is solved in the approximation that azimuthal asymmetries due to incoming polarized $e^\pm$ can be neglected.

Relative flavour composition:
It is possible to generate a single specified flavour, or to have the program generate the correct mixture of different flavours. The formulae given in section 2.1.1 are easily translated into a recipe for the relative flavour composition. If initial state radiation is included, this composition is an explicit function of the (reduced) hadronic CM energy.

Final state QED radiation:
JETSET contains no provisions for final state QED radiation or for initial–final state interference effects.

Perturbative QCD approaches:
Two alternatives are offered for the perturbative QCD picture, parton showers and matrix elements. They are described below. The parton shower option is the preferred one; see section 2.2.

Parton showers:
The main points of the JETSET showering algorithm are as follows.

- It is a leading log algorithm, of the improved, coherent kind, i.e. with angular ordering. The basic processes are branchings of the kinds $q \to q+g$, $g \to g+g$, and $g \to q+\bar{q}$.
- It can be used for an arbitrary initial $q\bar{q}$ pair, or, in fact for any one, two or three given entities (partons, hadrons, leptons, photons, etc.), although only a quark or gluon will initiate a shower. For e^+e^- annihilations, only the $q\bar{q}$ possibility is of interest, so the others will not be discussed here.
- The $q\bar{q}$ system may be given in any frame, but the evolution is carried out in the CM frame of the showering partons.
- Energy and momentum are conserved exactly at each step of the showering process.

- In the first branchings of the initial q and $\overline{q}$ partons, an additional rejection technique is used, so as to reproduce the lowest order differential three-jet cross-section.
- In subsequent branchings, angular ordering (coherence effects) is imposed.
- Gluon helicity effects, i.e. correlations between the production plane and the decay plane of a gluon, can be included.
- The first order α_S expression is used, with the Q^2 scale given by (an approximation to) the transverse momentum-squared of a branching. The default Λ_{QCD}, which should not be regarded as a proper $\Lambda_{\overline{MS}}$, is 0.4 GeV.
- The parton shower is by default cut off at a mass scale of 1 GeV.

Let us now proceed with a more detailed description; see also [8,2].

Kinematical variables. The evolution variable Q^2_{evol} of section 2.2.2 is associated with the mass-squared m^2 of the evolving parton, and hence the evolution variable is $t = \ln(m^2/\Lambda^2)$. Parton virtualities are therefore well defined during the evolution.

For the splitting variable z, which describes the sharing of energy and momentum between the two daughters of a branching, four slightly different options are available. In the default alternative, which is the only one to be described here, z is defined in terms of energy fractions in the CM frame of the showering parton system (with a modification that will appear later). Specifically, consider the process

$$e^+e^- \to (\gamma/Z^0)_0 \to q_1\overline{q}_2, \qquad q_1 \to q_3 g_4, \tag{128}$$

with indices as will be used in the following. Thus, in the branching $q_1 \to q_3 g_4$, z_1 is formally a Lorentz invariant,

$$z_1 = \frac{p_0 p_3}{p_0 p_1} = \frac{E_3}{E_1}, \tag{129}$$

but the virtual γ/Z^0 four-momentum p_0 acts as a 'gauge fixing' vector. The process is therefore not completely Lorentz covariant.

Cutoff of singularities. In order to cut off any singularities at $z = 0$ and $z = 1$, and to define a stopping point of the evolution in t, a parameter m_{min}, by default 1 GeV, is introduced. Partons are assumed to have effective masses

$$\begin{aligned} m_{eff,g} &= \frac{1}{2} m_{min}, \\ m_{eff,q} &= \sqrt{\frac{1}{4} m_{min}^2 + m_q^2}, \end{aligned} \tag{130}$$

where m_q are the ordinary quark masses in JETSET. For a parton to be able to branch at all, it must then have a mass larger than that of its lightest daughter pair,

$$\begin{aligned} m_{min,g} &= 2 m_{eff,g} = m_{min}, \\ m_{min,q} &= m_{eff,q} + m_{eff,g}. \end{aligned} \tag{131}$$

Evolution strategy. The evolution is performed using the standard Sudakov form factor approach, described in section 2.2.2. There is no pretabulation of form factors at an initialization stage, instead JETSET makes use of the 'veto algorithm'. The trick is

to find a simpler function, $\tilde{h}(t,z)$, everywhere larger than the correct integrand $h(t,z) = (\alpha_S(t,z)/2\pi) P_{a\to bc}(z)$, such that $\tilde{h}$ is analytically integrable, first over z (which becomes easier if also the z boundaries are given in a simplified form overestimating the correct range), and then over t, and such that the primitive function $\tilde{H}(t) = \int_{t_{min}}^{t} dt \int dz \tilde{h}(t,z)$ is analytically invertable. Starting from a known t_{max}, a t can then be found by using the modified Sudakov form factor $\tilde{S}(t) = \exp(-\tilde{H}(t))$. Since $\tilde{h}(t,z) \geq h(t,z)$ everywhere, a branching will, on the average, be selected 'too soon', i.e. with a too large t. In the veto algorithm, a selected (t,z) pair is accepted with a probability $h(t,z)/\tilde{h}(t,z)$. If rejected, the evolution is continued downwards, with the original t_{max} replaced by the just rejected t value. The process is iterated until a branching is accepted. It can be shown that, this way, the accepted branchings are distributed according to the correct Sudakov form factor.

At each stage of the evolution process, the evolution of a pair of partons is considered in parallel. At the first stage, it is the initial q and $\overline{q}$, at the second it is the two daughters of q and, separately, the two daughters of $\overline{q}$, etc. This pairwise arrangement is due to a coupling between the kinematical limits of the pair during the evolution process, and is a specific feature of the JETSET algorithm. As a consequence, it is possible to construct the parton four-momenta during the evolution process proper.

The evolution process. In the first step, partons 1 (q) and 2 ($\overline{q}$) are assigned masses m_1 and m_2, and simultaneously the z_1 and z_2 variables specifying the energy sharing between their respective daughters are determined. In the evolution of parton 1, it is assumed that the daughters 3 and 4 are not going to branch in their turn, and so they can be assigned the m_{eff} masses defined above. The allowed z range $z_{1-} < z_1 < z_{1+}$ is then

$$z_{1\pm} = \frac{1}{2}\left[1 \pm \beta_1 \theta(m_1 - m_{min,q})\right], \qquad \beta_1 = \frac{|\overline{p}_1|}{E_1}, \tag{132}$$

where $\theta(x)$ is the step function. Since the β_1 value depends on the mass of parton 2, this introduces the aforementioned coupling in the evolution of partons 1 and 2. The evolution of each initial parton is therefore started with a maximum allowed mass equal to the CM energy of the system, and evolved as if the other initial parton were massless. Once a pair of masses and z values have been chosen, it can be checked that $m_1 + m_2 < E_{CM}$, and that the z_1 and z_2 values are inside the physical range allowed by eq. (132) and its analog for parton 2, with the proper energies and momenta to use now known. If either z value is outside its allowed range, the evolution of the corresponding parton is continued downwards from the recently rejected mass; if both z values are rejected, the parton with largest mass is evolved further. The procedure is an extension of the veto algorithm mentioned earlier, where an initial overestimation of the allowed z range is compensated by rejection of some branchings. One should note, however, that the veto algorithm is not strictly applicable for the coupled evolution in two variables, and that therefore some arbitrariness is involved. This is manifest in the choice of which parton will be evolved further if both z values are unacceptable, or if $m_1 + m_2 > E_{CM}$.

When acceptable values have been found, the four-momenta p_1 and p_2 may be constructed. The angular orientation of that pair is assumed unchanged compared to the originally on-mass-shell $q\overline{q}$ pair. If the daughters q_3 and g_4 are assumed massless, then light-cone four-vectors $p_3^{(0)}$ and $p_4^{(0)}$ may be constructed from a knowledge of $\overline{p}_1$, m_1, z_1, and an isotropically selected azimuthal angle ϕ_1. The evolution of the daughters 3 and 4 may now be commenced, subject to the constraints $m_3 < E_3^{(0)} = z_1 E_1$,

$m_4 < E_4^{(0)} = (1-z_1)E_1$, and $m_3 + m_4 < m_1$. The allowed ranges $z_{3\pm}$ and $z_{4\pm}$ are defined in analogy with eq. (132). Again the coupled pairs of z_3, m_3 and z_4, m_4 values may be incompatible; in that case the parton with largest ratio m_{now}/m_{max} is evolved further, where m_{now} is the currently selected mass value and m_{max} the initially determined maximum mass.

When mass and z values have been fixed, corrected on-mass-shell four-vectors are calculated as

$$p_{3,4} = p_{3,4}^{(0)} \pm (r_4 p_4^{(0)} - r_3 p_3^{(0)}), \tag{133}$$

where

$$r_{3,4} = \frac{m_1^2 - \sqrt{(m_1^2 - m_3^2 - m_4^2)^2 - 4m_3^2 m_4^2} \pm (m_4^2 - m_3^2)}{2m_1^2}. \tag{134}$$

In other words, the meaning of the z_1 variable is somewhat reinterpreted, once the daughter masses are known. (The formulae are equivalent to the procedure previously used for partons 1 and 2, i.e. in the rest frame of 3 and 4 the angular orientation is unchanged).

The process can now be repeated for the decay of the parton pair produced by the original $\overline{q}$, and then for the parton pair produced by the branching of parton 3, etc. As the shower evolves, it becomes more and more likely that a given parton does not branch above its $m_{min,a}$ ($a = q$ or g) scale. These partons are assumed not to branch at all, and are put on mass-shell, i.e. $m = m_q$ for a quark and $m = 0$ for a gluon.

Matching to matrix elements. At the branchings of the two initial partons, an algorithm is used to match on to the first order three-jet matrix elements, as follows. A three-jet event $q(x_1)\overline{q}(x_2)g(x_3)$, with $x_i = 2E_i/E_{CM}$, can in the shower be obtained either by $\gamma/Z^0(0) \to q(1^*) + \overline{q}(2)$ followed by $q(1^*) \to q(1)g(3)$, or by a corresponding process via a $\overline{q}(2^*)$. In the first alternative the shower evolution variables are

$$\begin{aligned} m^2 = m_{1^*}^2 = (p_1 + p_3)^2 = (1-x_2)E_{CM}^2 \quad &\Rightarrow \quad dt = \frac{dm^2}{m^2} = \frac{dx_2}{1-x_2}, \\ z = \frac{p_0 p_1}{p_0 p_{1^*}} = \frac{x_1}{x_1 + x_3} = \frac{x_1}{2 - x_2} \quad &\Rightarrow \quad dz = \frac{dx_1}{2-x_2}. \end{aligned} \tag{135}$$

In the definitions of x_1 and x_2, masses other than m_{1^*} are neglected. Combined with the 2^* possibility, obtainable by an interchange of labels 1 and 2, the net branching probability given by the Altarelli-Parisi equations has the same singularity structure as the first order three-jet matrix element,

$$\begin{aligned} \frac{1}{\sigma}\frac{d\sigma}{dx_1 dx_2} &= \frac{4}{3}\frac{\alpha_S}{2\pi}\frac{A(x_1,x_2)}{(1-x_1)(1-x_2)}, \\ A_{shower}(x_1,x_2) &= 1 + \frac{1-x_1}{(1-x_1)+(1-x_2)}\left(\frac{x_1}{2-x_2}\right)^2 + \frac{1-x_2}{(1-x_1)+(1-x_2)}\left(\frac{x_2}{2-x_1}\right)^2, \\ A_{matrix}(x_1,x_2) &= x_1^2 + x_2^2. \end{aligned} \tag{136}$$

Here $A_{matrix}(x_1,x_2) \leq A_{shower}(x_1,x_2)$ everywhere; the ratio is never larger than 20/9 (equality for $x_1 = x_2 = 1/2$), and is unity in the limit x_1 (or x_2) $\to 1$. Branchings generated as part of the shower, and hence according to a weight proportional to A_{shower}, can therefore be accepted with a probability $A_{matrix}(x_1,x_2)/A_{shower}(x_1,x_2)$. This is made as part of the veto algorithm, i.e. in case of rejection evolution is continued downwards

from the rejected mass value. The procedure is used separately both for the initial q and the initial $\overline{q}$; in either case the other initial parton is assumed not to branch at all, so as to make x_1 and x_2 well-defined.

One should note that some implicit differences remain between the shower simulation of three-jets and the normal matrix element one. One is that the shower includes a Sudakov form factor to dampen the branching probability, and thus avoids the matrix element problem of probabilities potentially larger than unity. Another is that the α_S scale is given by the transverse momentum of branchings, see below.

Angular ordering. As the algorithm has been described so far, it provides a shower of the conventional branching kind. It is now necessary to impose decreasing emission angles. Since the initial angle between partons 1 and 2 is 180°, angular ordering gives no constraint on the angle θ_1 in the branching $1 \to 3 + 4$ — here the matching to the matrix element is instead at play. When parton 3 branches in its turn, however, the angle θ_3 should be smaller than θ_1, and correspondingly for subsequent branchings. The angular ordering is introduced as follows. The opening angle θ_a for a branching $a \to bc$ is approximately

$$\begin{aligned} \theta_a &\approx \frac{p_{Tb}}{E_b} + \frac{p_{Tc}}{E_c} \\ &\approx \sqrt{z_a(1-z_a)}m_a \left[\frac{1}{z_a E_a} + \frac{1}{(1-z_a)E_a}\right] \\ &= \frac{m_a}{\sqrt{z_a(1-z_a)}E_a}, \end{aligned} \tag{137}$$

so the requirement $\theta_3 < \theta_1$ is reduced to

$$\frac{z_3(1-z_3)}{m_3^2} > \frac{1-z_1}{z_1 m_1^2}, \tag{138}$$

where $E_3 = z_1 E_1$ has been used to eliminate the energy factors. If a branching of parton 3 does not fulfill the ordering condition, the branching is rejected and the evolution continued, i.e. the test is built into the veto algorithm framework. Eq. (138) is easily generalized to subsequent branchings.

Other shower aspects. A first order α_S is used, with argument $Q^2 = z(1-z)m^2 \approx p_T^2$, where p_T is the transverse momentum of a branching. The treatment of α_S at flavour thresholds is fairly primitive: only n_f is changed, without a corresponding redefinition of the Λ value. Since α_S is not defined for $z(1-z)m^2 \leq \Lambda^2$, a further constraint $z(z-1)m^2 > m_{min}^2/4$ is introduced on allowed branchings, where the requirement $m_{min} > 2\Lambda$ is automatically imposed by the program, if it is not already fulfilled. As alternatives, it is possible to have an $\alpha_S(m^2)$, or even a fixed α_S, in order to test the sensitivity of shower properties to the α_S choice.

Previously it was said that azimuthal angles in branchings were chosen isotropically. In fact, as an option, it is possible to include some effects of gluon polarization, which correlate the production and the decay planes of a gluon, such that a $g \to gg$ branching tends to take place in the production plane of the gluon, while a decay out of the plane is favoured for $g \to q\overline{q}$. The formulae are given e.g. in [14], as simple functions of the z value at the vertex where the gluon is produced and of the z value when it branches.

Finally, it should be noted that an Abelian vector toy model is included; in this option $g \to g + g$ branchings are absent, and $g \to q + \overline{q}$ ones enhanced (Casimir factors $C_F = 4/3 \to 1$, $N_C = 3 \to 0$, $T_R = n_f/2 \to 3n_f$).

Matrix elements:
The GKS second order matrix elements are available as an option to the parton shower default. The algorithm has already been described in section 3.2, to which the reader should turn for further details. Also within the matrix element option, a number of alternatives are available. The main ones include the generation

- of $q\overline{q}$ events only, i.e. the naive parton model;
- of $q\overline{q}$ and $q\overline{q}g$ events according to first order QCD;
- of $q\overline{q}g$, $q\overline{q}gg$ or $q\overline{q}q'\overline{q}'$ events only;
- of $q\overline{q}$ and $q\overline{q}g$ events according to a scalar gluon toy model; or
- of $q\overline{q}$, $q\overline{q}g$, $q\overline{q}gg$ and $q\overline{q}q'\overline{q}'$ events according to an Abelian vector toy model (in addition to a change in the four-jet sector, this option also implies large negative virtual corrections to the three-jet rate, but these latter can be switched off, if need be).

In addition one may use a fixed α_S or an α_S which runs according to the first or second order expression.

Angular orientation of events:
In the parton shower approach, the jet axis of the two original partons are chosen according to the lowest order electroweak formulae. Specifically, a polar angle θ is selected for the q, as described in section 2.1.2, including an energy dependent forward-backward asymmetry. The formulae used also include the effects of non-zero quark masses. If the incoming $e^{\pm}$ beams are unpolarized, the azimuthal angle ϕ is chosen isotropically, while transversely polarized beams gives rise to anisotropies. Once a jet axis has been chosen, the subsequent shower evolution treatment is assumed to be azimuthaily symmetric around this axis.

In the matrix element option, all mass corrections are neglected (unless only two-jet events are to be generated). For two-jet events, polar and azimuthal angles, θ and ϕ, are chosen as above. For three-jet events, it is necessary to select an additional angle χ, which describes the rotation of the three-jet event around its axis. In JETSET the quark defines the jet axis, as a matter of convention (the final result is, of course, independent of this choice). The orientation procedure is then:

1. start with the event being in the zx plane, with the quark along the $+z$ axis and the antiquark in the $+x$ half-plane;
2. rotate the event an angle χ around the z axis;
3. rotate the event an angle θ in the xz plane; and
4. rotate the event an angle ϕ around the z axis.

The necessary cross-sections are taken from [15], and depend on the x_1 and x_2 variables, i.e. the energy fractions of the quark and antiquark. The formulae are tedious, but straightforward to apply. They contain all Born level electroweak effects, including that of beam polarization.

No formulae are included for the angular orientation of four-jets. Instead the gg

pair of a $q\overline{q}gg$ event, or the $q'\overline{q}'$ pair of a $q\overline{q}q'\overline{q}'$ one, are joined into one effective gluon. The gluon mass is removed as in the $\overline{p}$ parton recombination scheme, so that afterwards effective three-jet variables x_1 and x_2 can be defined. After the angles χ, θ and ϕ have been chosen, as if the event were a three-jet one, the rotation procedure is applied to the actual four-jet configuration. Here initially the q and $\overline{q}$ are oriented as in the three-jet case while, of course, now the gg or $q'\overline{q}'$ pair is no longer in the plane.

Finally, if an initial state photon has been emitted, the event is boosted and rotated to take this into account, following the procedure suggested by Berends, Kleiss and Jadach [12]. Thus the photon is assigned to have been emitted either from the initial e^+ or e^-, according to kinematics-dependent relative probabilities, and transformations are defined between the e^+e^- CM frames before and after the photon emission, in both cases with the $e^{\mp}$ moving in the $\pm z$ direction in that frame. The partons are transformed from the e^+e^- after-photon-emission frame, i.e. the partonic CM frame, to the overall CM frame. The procedure for initial state photons is common for shower and matrix element treatments.

Fragmentation:
The main fragmentation option is the Lund string scheme, but independent fragmentation options are also available. The subsequent four sections give further details; the first one on flavour selection, which is common to the two approaches, the second on string fragmentation, the third on independent fragmentation, while the fourth and final contains information on a few other minor issues.

Flavour selection in the fragmentation:
In either string or independent fragmentation, an iterative approach is used to describe the fragmentation process. Given a quark q_1, it is assumed that a new $q_2\overline{q}_2$ pair may be created, such that a meson $q_1\overline{q}_2$ is formed, and a q_2 is left behind. This q_2 may at a later stage pair off with a $\overline{q}_3$, and so on. What need be given is thus the relative probabilities to produce the various possible $q_2\overline{q}_2$ pairs, $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, etc., and the relative probilities that a given $q_1\overline{q}_2$ quark pair combination forms a specific meson, e.g. for $u\overline{d}$ either π^+ or ρ^+.

In JETSET, it is assumed that the two aspects can be factorized, i.e. that it is possible first to select a $q_2\overline{q}_2$ pair, without any reference to allowed physical meson states, and that, once the $q_1\overline{q}_2$ flavour combination is given, it can be assigned to a given meson state with total probability unity.

Flavour composition. The production of $q'\overline{q}'$ pairs is assumed given by a quantum mechanical tunnelling process, as discussed in section 2.3.2. Inside the program, an effective parametrization is used, with the relative production of $q'\overline{q}'$ pairs given to be $u\overline{u} : d\overline{d} : s\overline{s} = 1 : 1 : \gamma_s$, where by default $\gamma_s = 0.3$. There is no production of heavier flavours in the fragmentation process, but only as part of the shower evolution.

Meson states. In the program, six meson multiplets are included. If the nonrelativistic classification scheme is used, i.e. mesons are assigned a valence quark spin S and an internal orbital angular momentum L, with the physical spin denoted J, $\overline{J} = \overline{L} + \overline{S}$, then the multiplets are:

- $L = 0$, $S = 0$, $J = 0$: the ordinary pseudoscalar meson multiplet;
- $L = 0$, $S = 1$, $J = 1$: the ordinary vector meson multiplet;

- $L = 1$, $S = 0$, $J = 1$: an axial vector meson multiplet;
- $L = 1$, $S = 1$, $J = 0$: the scalar meson meson multiplet;
- $L = 1$, $S = 1$, $J = 1$: another axial vector meson multiplet; and
- $L = 1$, $S = 1$, $J = 2$: the tensor meson multiplet.

Each multiplet has the full four-generation setup of 8×8 states included in the program, although many could never actually be produced. Some simplifications have been made; thus there is no mixing included between the two axial vector multiplets.

In the program, the spin S is first chosen to be either 0 or 1. This is done according to parametrized relative probabilities, where the probability for spin 1 is taken to be 0.5 for a meson consisting only of u and d quark, 0.6 for one which contains s as well, and 0.75 for quarks with c or heavier quark. The rationale is that one expects a ratio $1:3$ for spin $0:1$, from counting the number of states, but that the production of spin 1 states is suppressed because of the larger mass. Since the mass split is largest for the lighter mesons, it is also here that one could expect to find the $1:3$ ratio most violated.

By default, it is assumed that $L = 0$, such that only pseudoscalar and vector mesons are produced. For inclusion of $L = 1$ production, four parameters can be used, one to give the probability that a $S = 0$ state also has $L = 1$, the other three for the probability that a $S = 1$ state has $L = 1$ and J either 0, 1, or 2.

For the flavour diagonal meson states $u\bar{u}$, $d\bar{d}$ and $s\bar{s}$, it is also necessary to include mixing into the physical mesons. This is done according to a parametrization, based on the mixing angles given in the Review of Particle Properties [16].

Diquark mechanism for baryon production. Baryon production may, in its simplest form, be obtained by assuming that any flavour q_i given above could represent, either a quark, or an antidiquark in a colour triplet state. Then the same basic machinery can be run through as above, supplemented with the probability to produce various diquark pairs. In principle, there is one parameter for each diquark, but if tunnelling is still assumed to give an effective description, mass relations can be used to reduce the effective number of parameters. There are three main ones appearing in the program: the relative probability to pick a $\overline{qq}$ rather than a q, the extra suppression associated with a diquark containing a strange quark (over and above the ordinary s/u suppression factor γ_s), and the suppression of spin 1 diquarks relative to spin 0 ones (neglecting the factor of 3 difference from the number of spin states).

Baryon states. Only two baryon multiplets are included, i.e. there are no $L = 1$ excited states. The two multiplets are:

- $S = J = 1/2$: the 'octet' multiplet of SU(3) (in the full four-generation scenario in the program 168 states are available);
- $S = J = 3/2$: the 'decuplet' multiplet of SU(3) (120 states in the program).

In contrast to the meson case, different flavour combinations have different numbers of states available: for uuu only Δ^{++}, whereas uds may become either Λ, Σ^0 or Σ^{*0}.

An important constraint is that a baryon is a symmetric state of three quarks, neglecting the colour degree of freedom. When a diquark and a quark are joined to form a baryon, the combination is therefore weighted with the probability that they form a symmetric three-quark state. The program implementation of this principle is to first select a diquark at random, with the strangeness and spin 1 suppression factors

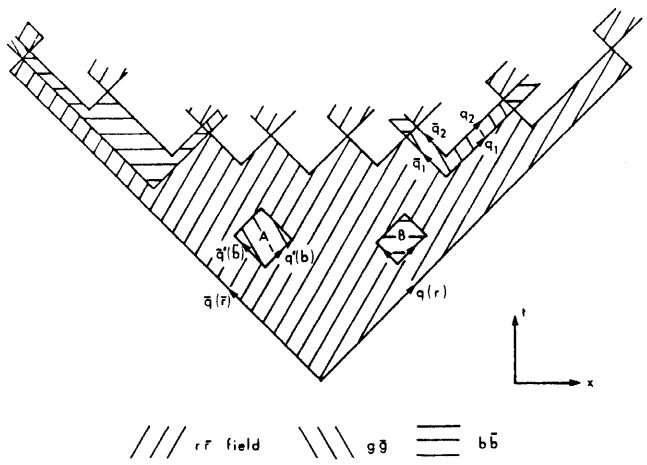

Figure 25: A colour field is stretched between a quark q and an antiquark $\bar{q}$. If virtual $q'\bar{q}'$ pairs are produced, the colour field can be changed from red-antired to green-antigreen or blue-antiblue. The direction of the field (going from triplet to antitriplet) is also changed. If a second pair is produced inside such a region, the colour field can break and a baryon-antibaryon pair can be produced (to the right). If two pairs are produced inside the region, a messon is formed between the baryon and the antibaryon (to the left).

above included, but then to accept the selected diquark with a weight proportional to the number of states available for the quark-diquark combination. In case of rejection, a new diquark is selected and tested, etc. A corresponding procedure is used for the quark selection when a diquark has already been formed in the previous step. Apart from the tunnelling suppression factors, and influence from the initial $q\bar{q}$ pair, all baryons are therefore represented according to their number of spin states, $2S + 1$.

Popcorn mechanism for baryon production. A more general framework for baryon production is the 'popcorn' one [6], in which diquarks as such are never produced, but rather baryons appear from the successive production of several $q'\bar{q}'$ pairs. The resulting space-time picture is illustrated in Fig. 25. The baryon-antibaryon pair need no longer be nearest neighbours, but can be formed with one or several mesons in between. In the program, only one intermediate meson is allowed. The relative proportion of the normal diquark mechanism (which may be considered a subprocess of the popcorn scenario) and the new popcorn one is an extra parameter; by default the two are assumed equally likely. Two additional parameters are related to strangeness suppression.

In total, the flavour iteration procedure therefore contains the following possible subprocesses (plus, of course, their charge conjugates):

- $q_1 \to q_2 + (q_1\bar{q}_2)$ meson;
- $q_1 \to \bar{q}_2\bar{q}_3 + (q_1q_2q_3)$ baryon;
- $q_1q_2 \to \bar{q}_3 + (q_1q_2q_3)$ baryon;

- $q_1 q_2 \to q_1 q_3 + (q_2 \overline{q}_3)$ meson;

with the constraint that the last process cannot be iterated to obtain several mesons in between the baryon and the antibaryon.

String fragmentation:
An iterative procedure can also be used for other aspects of the fragmentation. This is possible because, in the string picture, the various points where the string break by the production of $q\overline{q}$ pairs are causally disconnected. Whereas the space-time picture in the CM frame is such that slow particles (in the middle of the system) are formed first, this ordering is Lorentz frame dependent and hence irrelevant. One may therefore make the convenient choice of starting an iteration process at the ends of the string and proceeding towards the middle.

The string fragmentation scheme is rather complicated for a generic multiparton state. In order to simplify the discussion, we will therefore start with the simple $q\overline{q}$ process, and only later survey the complications that appear when additional gluons are present. (This distinction is made for pedagogical reasons, in the program there is only one general-purpose algorithm).

The Lund symmetric fragmentation function. Assume a $q\overline{q}$ jet system, in its CM frame, with the quark moving out in the $+z$ direction and the antiquark in the $-z$ one. We have discussed how it is possible to start the flavour iteration from the q end, i.e. pick a $q_1\overline{q}_1$ pair, form a hadron $q\overline{q}_1$, etc. In section 2.3.2 it has also been noted that the tunnelling mechanism is assumed to give a transverse momentum $\overline{p}_T$ for each new $q'\overline{q}'$ pair created, with the $\overline{p}_T$ locally compensated between the q' and the $\overline{q}'$ member of the pair, and with a Gaussian distribution in p_x and p_y separately. In the program, this is regulated by one parameter, which gives the root-mean-square p_T of a quark. Hadron transverse momenta are obtained as the sum of $\overline{p}_T$:s of the constituent q_i and $\overline{q}_{i+1}$, where a diquark is considered just as a single quark.

What remains to be determined is the energy and longitudinal momentum of the hadron. In fact, only one variable can be selected independently, since the momentum of the hadron is constrained by the already determined hadron transverse mass m_T,

$$(E + p_z)(E - p_z) = E^2 - p_z^2 = m_T^2 = m^2 + p_x^2 + p_y^2. \tag{139}$$

In an iteration from the quark end, one is led (by the desire for longitudinal boost invariance and other considerations) to select the z variable as the fraction of $E + p_z$ taken by the hadron, out of the available $E+p_z$. As hadrons are split off, the $E+p_z$ (and $E - p_z$) left for subsequent steps is reduced accordingly. The fragmentation function $f(z)$, which expresses the probability that a given z is picked, could in principle be arbitrary — indeed, several such choices can be used inside the program, including the ones mentioned in section 2.3.1.

If one, in addition, requires that the fragmentation process as a whole should look the same, irrespectively of whether the iterative procedure is performed from the q end or the $\overline{q}$ one, 'left-right symmetry', the choice is essentially unique,

$$f(z) \propto \frac{1}{z} z^{a_\alpha} \left(\frac{1-z}{z}\right)^{a_\beta} \exp\left(-\frac{b m_T^2}{z}\right). \tag{140}$$

There is one parameter a_i for each flavour i, with the index α corresponding to the 'old' flavour in the iteration process, and β to the 'new' flavour. It is customary to put all a_i

the same, and thus arrive at the simplified expression given in eq. (30). In the program, only two separate a values can be given, that for quark pair production and that for diquark one; by default the two are taken to be the same. In addition, there is the b parameter, which is universal.

Joining the jets. The $f(z)$ formula above is only valid, for a the breakup of a jet system into a hadron plus a remainder-system, when the remainder mass is large. If the fragmentation algorithm were to be used all the way from the q end to the $\overline{q}$ one, the mass of the last hadron to be formed at the $\overline{q}$ end would be completely constrained by energy and momentum conservation, and could not be on its mass shell.

The practical solution to this problem is to carry out the fragmentation both from the q and the $\overline{q}$ end, such that for each new step in the fragmentation process, a random choice is made as to from what side the step is to be taken. If the step is on the q side, then z is interpreted as fraction of the remaining $E + p_z$ of the system, while z is interpreted as $E - p_z$ fraction for a step from the $\overline{q}$ end. At some point, when the remaining mass of the system has dropped below a given value, it is decided that the next breakup will produce two final hadrons, rather than a hadron and a remainder-system. Since the momenta of two hadrons are to be selected, rather than that of one only, there is enough freedom to have both total energy and total momentum completely conserved.

The mass at which the normal fragmentation process is stopped and the final two hadrons formed is not actually a free parameter of the model: it is given by the requirement that the string everywhere look the same, i.e. that the rapidity spacing of the final two hadrons, internally and with respect to surrounding hadrons, be the same as elsewhere in the fragmentation process. The stopping mass, for a given setup of fragmentation parameters, has therefore been determined in separate runs. If the fragmentation parameters are changed, some retuning should be done but, in practice, reasonable changes can be made without any special arrangements. The basic stopping mass value is modified according to the masses of the quarks involved, see [1,2] for details.

If the remainder-mass is large enough, there are two kinematically allowed solutions for the final two hadrons (two mirror images in the rest frame of the remainder-system). The choice between these two solutions is also given by the consistency requirements, and is suitably parametrized in the program. If, on the other hand, the transverse mass of the remainder-system is smaller than the sum of transverse masses of the final two hadrons, the whole fragmentation chain is rejected, and started over from the q and $\overline{q}$ endpoints. The same happens if the very final hadron, for which the flavour content is completely constrained, happens to be a diquark-antidiquark state or, in the popcorn mechanism, contains a diquark which was supposed to split into a meson plus a new diquark.

While some practical compromises have to be accepted in the joining procedure, the fact that the joining takes place in different parts of the string in different events means that, in the end, essentially no visible effects remain.

String motion and infrared stability. In moving on to the fragmentation of multiparton states, it is worthwhile to first consider the string motion in a a three-jet $q\overline{q}g$ event. Such an event initially corresponds to having a string stretched from the q via the g to the $\overline{q}$, i.e. two string pieces. In the string piece between the g and the q ($\overline{q}$), g

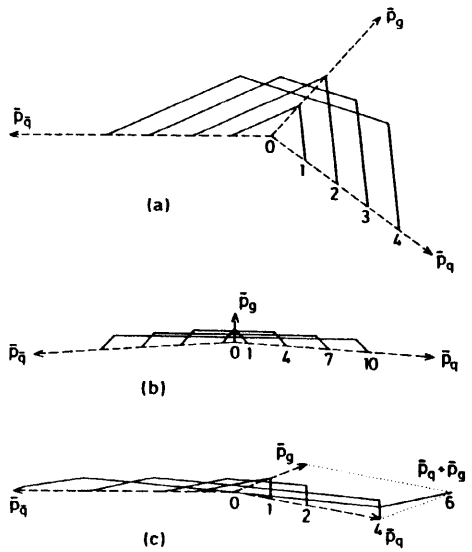

Figure 26: The string drawing for (a) an ordinary three-jet event, (b) a three-jet event with a soft gluon, and (c) a three-jet event with a collinear gluon. Dashed lines give the momenta (and hence the trajectories) of the partons. Full lines give the string shape at different times, with numbers representing time in some suitable scale.

four-momentum is flowing towards the q ($\overline{q}$) end and q ($\overline{q}$) four-momentum towards the g end. Such packets of energy and momentum are called 'genes' [17]. When the gluon has lost all its energy, the g four-momentum continues moving away from the middle (i.e. where the gluon used to be), and instead a third string region is formed there, consisting of inflowing q and $\overline{q}$ four-momentum, Fig. 26a. If this third region would only appear at a time later than the typical time scale for fragmentation, it could not affect the sharing of energy between different particles. This is true in the limit of high energy, well separated partons.

For a small gluon energy, on the other hand, the third string region appears early, and the overall drawing of the string becomes fairly two-jetlike, Fig. 26b. In the limit of vanishing gluon energy, the two initial string regions collapse to naught, and the ordinary two-jet event is recovered. Also for a collinear gluon, i.e. θ_{qg} (or $\theta_{\overline{q}g}$) small, the stretching becomes two-jetlike, Fig. 26c. These properties of the string motion are the reason why the string fragmentation scheme is 'infrared safe' with respect to soft or collinear gluon emission.

Fragmentation of multiparton systems. The full machinery needed for a multiparton system is very complicated, and is described in detail in [5]. The following outline is far from complete, and is complicated nonetheless. The main message to be conveyed is that a Lorentz covariant algorithm exists for handling an arbitrary parton configuration, but that the necessary machinery is more complex than in either cluster or independent fragmentation.

Assume n partons, with ordering along the string, and related four-momenta, given by $q(p_1)g(p_2)g(p_3)\cdots g(p_{n-1})\overline{q}(p_n)$. The initial string then contains $n-1$ separate pieces. The string piece between the quark and its neigbouring gluon is, in four-momentum space, spanned by one side with four-momentum $p_+^{(1)} = p_1$ and another with $p_-^{(1)} = p_2/2$.

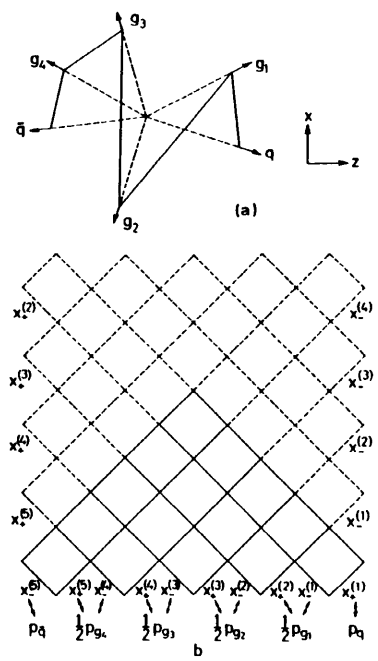

Figure 27: Representation of an arbitrary jet system, in this case $qg_1g_2g_3g_4\overline{q}$. (a) Space picture of the string configuration before any hadron has lost its energy. (b) Parameter space representation of the 41 string regions that exist over half a period of the full string motion. The five initial regions, bottom row, and the other ten actually accessible in the fragmentation algorithm are shown in full. Bottom line shows how the x variables are related to the parton momenta.

The factor of 1/2 in the second expression comes from the fact that the gluon shares its energy between two string pieces. The indices '+' and '-' denotes direction towards the q and $\overline{q}$ end. The next string piece, counted from the quark end, is spanned by $p_+^{(2)} = p_2/2$ and $p_-^{(2)} = p_3/2$, and so on, with the last one being $p_+^{(n-1)} = p_{n-1}/2$ and $p_-^{(n-1)} = p_n$.

For the algorithm to work, it is important that all $p_\pm^{(i)}$ be light-conelike, i.e. $p_\pm^{(i)2} = 0$. Since gluons are massless, it is only the two endpoint quarks which can cause problems. The procedure here is to create new $p_\pm$ vectors for each of the two endpoint regions, defined to be linear combinations of the old $p_\pm$ ones for the same region, with coefficients determined so that the new vectors are light-conelike. De facto, this corresponds to replacing a massive quark at the end of a string piece with a massless quark at the end of a somewhat longer string piece. With the exception of the added fictitious piece, which anyway ends up entirely within the heavy hadron produced from the heavy quark, the string motion remains unchanged by this.

In the continued string motion, when new string regions appear as time goes by, cf. Fig. 26, each such string region can be represented as being spanned by one $p_\pm^{(j)}$ and another $p_\pm^{(k)}$ four-vector. The classification scheme is illustrated in Fig. 27. The regions made up of two p_+ or two p_- momenta appear when an endpoint quark has lost all its original momentum, has accreted the momentum of an gluon, and is now reemitting this momentum. In practice, these regions may be neglected, so that only the string regions shown in full need be considered. These are all made up by a $(p_+^{(j)}, p_-^{(k)})$ pair of momenta.

In Fig. 27b, the fragmentation procedure can be described as a sequence of steps, starting at the rightmost corner of the bottom row line of squares (the quark end) and ending at the leftmost corner of the bottom row (the antiquark end). Each step is taken from an 'old' $q'\overline{q}'$ pair production vertex, to the production vertex of a 'new' $q'\overline{q}'$ pair, and the string piece between these two string breaks represent a hadron. The four-momentum of each hadron can be read off, for p_+ (p_-) momenta, by projecting the separation between the old and the new vertex along the diagonal downwards right (downwards left). If the four-momentum fraction of $p_\pm^{(i)}$ taken by a hadron is denoted $x_\pm^{(i)}$, then the total hadron four-momentum is given by

$$p = \sum_{j=j_1}^{j_2} x_+^{(j)} p_+^{(j)} + \sum_{k=k_1}^{k_2} x_-^{(k)} p_-^{(k)} + p_{x1} e_x^{(j_1 k_1)} + p_{y1} e_y^{(j_1 k_1)} + p_{x2} e_x^{(j_2 k_2)} + p_{y2} e_y^{(j_2 k_2)}, \qquad (141)$$

for a step from region (j_1, k_1) to region (j_2, k_2). By necessity, $x_+^{(j)}$ is unity for a $j_1 < j < j_2$, and correspondingly for $x_-^{(k)}$. The (p_x, p_y) pairs are the transverse momenta produced at the two string breaks, and the (e_x, e_y) pairs four-vectors transverse to the string directions in the regions of the respective string breaks:

$$\begin{aligned} e_x^{(jk)2} = e_y^{(jk)2} &= -1, \\ e_x^{(jk)} e_y^{(jk)} = e_{x,y}^{(jk)} p_+^{(j)} = e_{x,y}^{(jk)} p_-^{(k)} &= 0. \end{aligned} \qquad (142)$$

The fact that the hadron should be on mass-shell, $p^2 = m^2$, puts one constraint on where a new breakup may be, given that the old one is already known, just as eq. (139) did in the simple two-jet case. The other one is, as before, to be given by the fragmentation function $f(z)$. The interpretation of such a z is only well-defined for a step entirely constrained to one of the initial string regions, however, which is not enough. In the two-jet case, the z values can be related to the invariant times of string breaks, as follows. The variable $\Gamma = (\kappa\tau)^2$, with κ the string tension and τ the invariant time between the production vertex of the partons and the breakup point, obeys an iterative relation of the kind

$$\begin{aligned} \Gamma_0 &= 0, \\ \Gamma_i &= (1 - z_i)\left(\Gamma_{i-1} + \frac{m_{Ti}^2}{z_i}\right). \end{aligned} \qquad (143)$$

Here Γ_0 represents the value at the q and $\overline{q}$ endpoints, and Γ_{i-1} and Γ_i the values at the old and new breakup vertices needed to produce a hadron with transverse mass m_{Ti}, and with the z_i of the step chosen according to $f(z_i)$. The invariant time can be defined in an unambiguous way, also over boundaries between the different string regions, so for multijet events the z variable may be interpreted just as an auxiliary variable needed to determine the next Γ value. (In the Lund symmetric fragmentation function derivation, the Γ variable actually does appear naturally, so the choice is not as arbitrary as it may seem here.) The mass and Γ constraints together are sufficient to determine where the next string breakup is to be chosen, given the preceding one in the iteration scheme. Actually, several ambiguities remain, but are of no importance for the overall picture.

The algorithm for finding the next breakup then works something like follows. Pick a hadron, p_T, and z, and calculate the next Γ. If the old breakup is in the region (jk), and if the new breakup is also assumed to be in the same region, then the m^2 and Γ

constraints can be reformulated in terms of the fractions $x_+^{(j)}$ and $x_-^{(k)}$ the hadron must take of the total four-vectors $p_+^{(j)}$ and $p_-^{(k)}$:

$$\begin{aligned} m^2 &= c_1 + c_2 x_+^{(j)} + c_3 x_-^{(k)} + c_4 x_+^{(j)} x_-^{(k)}, \\ \Gamma &= d_1 + d_2 x_+^{(j)} + d_3 x_-^{(k)} + d_4 x_+^{(j)} x_-^{(k)}. \end{aligned} \qquad (144)$$

Here the coefficients c_n are fairly simple expressions, obtainable by squaring eq. (141), while d_n are slightly more complicated in that they depend on the position of the old string break, but both the c_n and the d_n are explicitly calculable. What remains is an equation system with two unknowns, $x_+^{(j)}$ and $x_-^{(k)}$. The absence of any quadratic terms is due to the fact that all $p_\pm^{(i)} = 0$. Of the two possible solutions to the equation system (elimination of one variable gives a second degree equation in the other), one is unphysical and can be discarded outright. The other solution is checked for whether the $x_\pm$ values are actually inside the physically allowed region, i.e. whether the $x_\pm$ values of the current step, plus whatever has already been used up in previous steps, are less than unity. If yes, a solution has been found. If no, it is because the breakup could not take place inside the region studied, i.e. because the equation system was solved for the wrong region. One therefore has to change either index j or index k above by one step, i.e. go to the next nearest string region. In this new region, a new equation system of the type in eq. (144) may be written down, with new coefficients. A new solution is found and tested, and so on until a physically acceptable solution is found. The hadron four-momentum is now given by an expression of the type (141). The breakup found forms the starting point for the new step in the fragmentation chain, and so on. The final joining in the middle is done as in the two-jet case, with minor extensions.

Independent fragmentation:
The independent fragmentation framework, available as an option in JETSET, is pretty much in agreement with what is discussed in section 2.3.1. Therefore only some minor points are given here.

Fragmentation schemes. Fragmentation functions can be chosen among those listed in section 2.3.1, but also here the default is the Lund symmetric fragmentation function.

Gluons may be treated either as a single quark jet, or as a $q\overline{q}$ jet pair, in which the q and $\overline{q}$ are parallel and share the gluon energy according to an Altarelli-Parisi type ansatz. Fragmentation functions and transverse momenta may be selected independently from the ordinary quark jet values.

Conservation issues. Flavour conservation can be achieved by an algorithm which leaves particle three-momenta unchanged, but reassigns flavours of centrally produced particles, in some minimal fashion.

Four schemes are available for energy and momentum conservation, in addition to the default one of having no conservation at all (except in an average sense).

- Conserve transverse momentum locally within each jet, such that energy and momentum conservation can be achieved by a rescaling of longitudinal momenta, separately for each jet.
- Boost the generated event so that net momentum vanishes, and afterwards rescale all three-momenta by a common factor, selected to give energy conservation.
- Shift all particle three-momenta by a common amount to give net vanishing mo-

mentum, and then rescale as before.

- Shift all particle three-momenta, for each particle by an amount proportional to the longitudinal mass with respect to the imbalance direction, and with overall magnitude selected to give momentum conservation, and then rescale as before.

In addition, there is a choice of whether to treat separate colour singlets (like $q\overline{q}'$ and $q'\overline{q}$ in a $q\overline{q}q'\overline{q}'$ event) separately or as one single big system.

Other fragmentation aspects:
Here two aspects are considered, which are applicable regardless of whether string or independent fragmentation is used.

Small mass systems. Occasionally, a jet system may have too small an invariant mass for the ordinary jet fragmentation schemes. This is particularly a problem when showers are used, since two nearby $g \to q' + \overline{q}'$ branchings may give rise to a low-mass colour singlet system. Before the ordinary fragmentation, there is included an optional additional step, to catch situations of this kind. First the jet system with lowest invariant mass, minus endpoint quark masses, is found. If this is too low for jet fragmentation, an attempt is made to split the system into two hadrons by producing a new $q'\overline{q}'$ pair (with q' chosen according to the standard fragmentation scheme, so that e.g. also diquarks are allowed) to go with the existing endpoint flavours. If the sum of the two hadron masses is smaller than the total invariant mass, a simple isotropic two-particle decay is performed. If not, the endpoint flavours are combined to give one single hadron. Next, the parton (or hadron) is found which, when taken together with the jet system, has the largest invariant mass. A minimal transfer of four-momentum is then performed, which puts the hadron on mass-shell while keeping the mass of the parton unchanged. With this done, one may again search for a low-mass jet system, and iterate the procedure above, if need be. The procedure may be seen as a 'poor man's cluster fragmentation', i.e. a cluster and a low-mass string are considered to be more or less the same thing.

Bose-Einstein effects. A crude option for the simulation of Bose-Einstein effects is included, but is turned off by default. Here the detailed physics is not that well understood, see e.g. [18]. What is offered is an algorithm, more than just a parametrization (since very specific assumptions and choices have been made), and yet less than a true model (since the underlying physics picture is rather fuzzy). In this scheme, the fragmentation is allowed to proceed as usual, and so is the decay of short-lived particles like ρ. Then pairs of identical particles, π^+ say, are considered one by one. The Q value is evaluated,

$$Q_{12} = \sqrt{(p_1 + p_2)^2 - (m_1 + m_2)^2}, \tag{145}$$

and a shifted (smaller) Q' is found such that the (infinite statistics) ratio $C_2(Q)$ of shifted to unshifted Q distributions is given by the requested parametrization. The shape may be chosen either exponential or Gaussian,

$$C_2(Q) = 1 + \lambda \exp\left(-(Q/d)^r\right), \qquad r = 1 \text{ or } 2. \tag{146}$$

(In fact, the distribution dips slightly below unity at Q values outside the Bose enhancement region, from conservation of total multiplicity.) The change of Q can be translated into an effective shift of the momenta of the two particles, if one uses as

extra constraint that the total three-momentum of each pair be conserved in the CM frame of the event. Only after all pairwise momentum shifts have been evaluated, with respect to the original momenta, are these momenta actually shifted, for each particle by the sum of evaluated shifts. The total energy of the event is slightly reduced in the process, which is compensated by an overall rescaling of all CM frame momentum vectors. Finally, the decay chain is resumed with more long-lived particles like π^0.

Two comments can be made. The Bose-Einstein effect is here interpreted almost as a classical force acting on the 'final state', rather than as a quantum mechanical phenomenon on the production amplitude. This is not a credo, but just an ansatz to make things manageable. Also, since only pairwise interactions are considered, the effects associated with three or more nearby particles tend to get overestimated. (More exact, but also more time-consuming methods may be found in [19].) Thus the input λ may have to be chosen smaller than what one wants to get out. This option should therefore be used with caution, and only as a first approximation to what Bose-Einstein effects can mean.

Particle masses:
Known particle masses are taken from the Review of Particle Properties [16]. Some charm and most heavier hadron masses are calculated according to mass formulae of the type

$$m = m_0 + \sum_i m_i + k \sum_{i<j} \frac{\langle \overline{\sigma}_i \overline{\sigma}_j \rangle}{m_i m_j}. \tag{147}$$

Here the m_i are the valence quark constituent masses, the $\langle \overline{\sigma}_i \overline{\sigma}_j \rangle$ are the quark spin-spin correlation expectation values, and m_0 and k are free parameters, determined for each multiplet separately, from the known spectroscopy.

Particles may also have a finite width. The mass is then selected according to a simple Breit-Wigner shape, either in m or in m^2, according to choice, but truncated, $|m - m_{nominal}| \leq \delta$. The δ value is chosen individually for each particle, such that $m_{nominal} - \delta$ is larger than the sum of (nominal) decay product masses for all possible decay channels. This crude approach is introduced so that the choice of a particle mass can be decoupled from the subsequent selection of a decay channel.

Particle decays:
Where known, decay channels and branching ratios are taken from the Review of Particle Properties [16]. For τ, D^0 and D^+, this information is complemented by additional guessed-at decay modes and branching ratios. Also a few other explicit branching ratios are based on guesswork.

For D_s^+, charmed baryons and all bottom hadrons, the more or less known semileptonic branching ratios are supplemented by phase-space models for the hadronic decays. Here an initial flavour content is given, e.g. $B_u^+ = u\overline{b} \to u\overline{c}u\overline{d}$, according to a reasonable composition (taking into account Cabibbo suppressed flavour combinations and annihilation graphs, where present). A hadron multiplicity is picked, according to a simple Gaussian parametrization, which depends on the mass of the decaying particle and the decay product flavour composition. Starting from the existing flavours, new $q'\overline{q}'$ pairs are created, according to exactly the same algorithm as used in the fragmentation (thus also diquark production is allowed) and paired off with the old flavours, or with each other, to give the desired number of hadrons. If the sum of hadron masses exceeds the

mother particle mass, the decay channel is rejected, and a new one is selected, starting with the same flavour content but a new hadron multiplicity. If the channel is accepted, particles are distributed according to phase-space.

For top hadrons, and other heavy objects, the decay is assumed to take place into a new partonic state. This state may contain one or two $q_1\overline{q}_2$ colour singlet systems. For a system made up out of a spectator quark and another quark, there is assumed to be no QCD shower evolution, since the mass of this state normally is small anyway. Special care is also taken to check if this system should be collapsed into one or two particles. For other systems, i.e. normally W decay products, shower evolution is performed. Thereafter jet systems are fragmented, and so on.

Most decays are performed according to phase-space, with up to ten decay products possible. Some exceptions exist, where matrix element information is included. The list contains

- ω or $\phi \to \pi^+\pi^-\pi^0$;
- π^0 or $\eta \to \gamma e^+e^-$ (Dalitz decays);
- the decay of a pseudoscalar into three pseudoscalars via an intermediate vector+pseudoscalar state;
- $\Upsilon \to ggg$ or γgg; and, of course,
- $V-A$ weak decay matrix elements, for top and heavier with W propagator effects included (and with special care taken so that the W mass generation is done efficiently, also if the W and t masses are more or less degenerate).

Utilities:
JETSET contains utility routines for:

- event listing, particle data listing and parameter listing;
- rotation and boosting of events;
- selective editing of the event record;
- editing of particle data (including masses, decay modes, and branching ratios);
- calculation of particle properties not stored (like, e.g., rapidity);
- sphericity analysis (also with linearized momentum dependence);
- thrust analysis;
- cluster finding, either according to a native version with a distance measure of a transverse momentum type, or according to the JADE invariant mass recipe [20] (which is the better of the two depends on what the reconstructed jets are to be used for);
- calorimeter cell cluster finding of the generic UA1 type (i.e. not directed towards e^+e^- applications);
- Fox-Wolfram moments analysis;
- statistics on parton and particle content of event (cumulative during course of run);
- statistics on factorial moments of the charged multiplicity distribution (cumulative); and
- statistics on the energy-energy correlation and its asymmetry (cumulative).

4.8.4 Some Comments on Program Limitations

JETSET is a model for hadron production, not a theory. It is therefore clear that some ideas work less well than others. On the other hand, JETSET is no worse off in this respect than are any of the other program presented in this report. Below, some of the limitations and unresolved issues are mentioned. In addition, the subsequent section contains a list of disagreements between program and data.

Electroweak physics. The electroweak cross-section formulae and the initial state radiation algorithm give a satisfactory description up to PETRA energies, and a passable description up to TRISTAN, but are not suited for precision studies at the Z^0 peak. In particular, it is known that the first order unexponentiated QED radiation formulae tend to overestimate the effects of photon emission on the naive Z^0 line shape. Further, the lack of electroweak loop effects in JETSET means that, even disregarding initial state radiation, the cross-section is incorrectly described in the Z^0 region. A number of more sophisticated generators for initial state radiation and hard interactions have been developed; these are described in the electroweak generators report. Due to the modular structure of JETSET, it should be easy to take a $q\overline{q}$ pair thus generated and feed it into JETSET for subsequent showering, fragmentation and decays.

The flavour composition formulae are reasonably complete. The behaviour at flavour thresholds is only crudely described, but this should not be important for LEP physics. The additional contributions to the $b\overline{b}$ cross-section that will appear for a heavy top, due to loop diagrams, are not included. If JETSET is to be interfaced with an initial state radiation program, it may make sense to let that program select the flavour as well, since that generally gives more flexibility in the interface structure.

Like all other programs, JETSET does not address the issues of final state QED radiation or initial–final state interference.

The parton shower. The JETSET shower algorithm is less well founded in theory than an algorithm like the HERWIG one. In particular, the choice of z variable and evolution strategy leads to somewhat different boundaries on the allowed z range for branchings than in HERWIG. One peculiar feature is that the evolution of two sister partons has to be considered in parallel, with interrelated kinematics cuts. All such correlations are absent in the HERWIG algorithm. While the differences formally should be of subleading character, this does not mean they need be small. Indeed, there are already some indications of such differences, e.g. in that the average parton multiplicity increase with CM energy is somewhat slower in JETSET than in HERWIG. It is here believed that the HERWIG algorithm does give the correct subleading behaviour.

Another difference is that the JETSET evolution is performed in terms of masses, with angular ordering imposed by a veto on nonordered branchings, while HERWIG has angular ordering built in already by the choice of evolution variable. Even apart from the z differences noted above, this means that there is a difference in the Sudakov form factor dampening. In JETSET, with its ordering in masses, large masses for daughters are preferred, since then the Sudakov form factor is close to unity; by the same token the HERWIG algorithm favours large emission angles, close to the maximum allowed one. Of course, one should not overinterpret the word 'favour'; in either case the regions of large Sudakov form factors also correspond to regions of small differential cross-section: the singularities of the emission process are at low mass and small angles.

The choice of z definition, discussed above, was selected so as to make the matching to first order matrix elements easier. This matching has proved phenomenologically successful, in providing a reasonable amount of hard gluon emission. It is still fairly unclear whether this is really a consequence of the shower matching procedure in itself; equally likely it could be a consequence of the z choice. Also, a proper amount of hard gluon radiation is no guarantee that the algorithm handles soft and collinear gluon emission properly, and obviously the latter ones are those which will determine the high energy behaviour of the algorithm.

Like most other algorithms presented here, the JETSET algorithm is not fully Lorentz covariant. Everything is in principle formulated in term of four-vectors and Lorentz scalars, but it is necessary to include the four-momentum vector p_0 of the initial γ/Z^0 in the construction of the invariants. This might seem reasonable for the branchings of the initial q and $\bar{q}$, but in the standard JETSET algorithm p_0 has to be used for the z definition in every step of the shower, however removed from the primary vertex. An alternative z definition exists in the program, where the reference vector is always related to the 'grandmother' parton of a branching, and hence is more locally defined, but in this scheme the allowed z range becomes too constrained to give a reasonable amount of shower evolution.

Finally, a note on the angular ordering conditions of eqs. (137) and (138). It might seem that the approximation used for the angle is fairly crude. One should remember, however, that angles are not Lorentz covariant objects, and so the statement of angular ordering is in itself somewhat fuzzy. In addition, the original derivation of coherence actually gave a constraint of the kind in eq. (138), which only afterwards was interpreted in terms of angular ordering.

Matrix elements. It has already been pointed out, in section 3.2, that the GKS second order matrix elements implemented in JETSET are incomplete. When comparing with data, this will give somewhat different α_S values than would be obtained with other matrix elements. Despite this, the JETSET matrix element package is elaborate, flexible and user-friendly. The lack of mass effects and the simplified angular orientation scheme for four-jets may be seen as shortcomings, but are still of minor importance for most LEP applications.

Fragmentation. Of the two fragmentation options, only string fragmentation should be taken seriously. The independent fragmentation algorithm, in itself, is as powerful as any presented in these proceedings, but we know that it fails to describe a number of experimental observations, and that there are a number of conceptual shortcomings.

The string fragmentation algorithm has been very successful experimentally. Certainly a number of compromises have been made in the practical implementation, like in the procedure to join the two fragmentation chains in the middle, in some details of the string motion, in the translation of the fragmentation function z into the step to be taken, and in other issues, but none of these are known to give any real problems.

The flavour production properties rest on shakier ground. Maybe the whole tunnelling concept is wrong. Certainly the neglect of hadron mass effects in flavour selection is incorrect, or at least only a crude approximation; else it would not have been necessary to include separate parameters for the vector to pseudoscalar production ratio according to the quark content. The baryon production scenario, with its popcorn mechanism, also seems to have problems in describing the data, despite a large number

of free parameters.

Particle decays. For very light and very heavy particles, JETSET should have a fairly realistic description of decays. It may be the intermediate range which contains the most problems: the neglect of τ polarization effects, the incomplete usage of experimental information on D^0 and D^+ decays, and the uncertain quality of the phase-space decay scheme used e.g. for B mesons, to mention the three most obvious cases.

4.8.5 Comparison with Data

For the last five years or so, JETSET has been the most frequently used e^+e^- QCD event generator for applications above the Υ region, both at PETRA, PEP and TRISTAN, and for LEP and SLC preparatory studies. Considering the verdict of the experimental groups, it seems that Lund string fragmentation still is the most successful fragmentation model we have today. As for the showering algorithm, it does not seem to fare worse than any other currently used. Much of these studies were carried out with version 6.3, which, from a physics point of view, is just about equivalent to 7.1. A few examples of comparison between model and data are given in section 5 of this report. A slightly longer discussion, from the horizon of the JETSET author, may be found in [21].

The main problems known in comparisons between data and model include.

- The B meson fragmentation spectrum is too hard. In principle, with the a and b parameters of the Lund symmetric fragmentation function determined by inclusive event properties, a unique prediction exist for charm and bottom fragmentation functions. Charm seems to come out reasonably well, but not bottom. Conceivably the problem could be, at least partly, in an underestimation of gluon emission from b quarks, but it could also signal a breakdown of the 'left-right' symmetry argument for leading particles.
- Despite a large number of flavour composition parameters, it is not possible to achieve a completely satisfactory description of particle production rates. In particular, the Ω^- rate seems to be underestimated by more than a factor 10 in the program as compared to PETRA/PEP data (whereas the difference with CLEO and ARGUS data seems to be less).
- The program overestimates the fraction of protons among high-momentum charged particles. This could signal the need for a separate a parameter for baryon production, but could also arise from more fundamental problems.
- Baryon-antibaryon correlations favour a low pure diquark pair production rate, i.e. most event seems to have (at least) one meson produced between the baryon and the antibaryon. It is then not a good approximation to only allow for one intermediate meson in the popcorn scenario, as is done in the program. Ultimately, the whole baryon production philosophy may be put in question.
- Raising the sight beyond the e^+e^- horizon, the model has met with additional problems. While most of these fall outside the frame of this report, it may be interesting to note that even a parameter as fundamental as the s/u quark suppression factor, γ_s, seems to vary from process to process, and with CM energy. This could signal a breakdown of jet universality, or serious model deficiencies.

4.8.6 Installation and Running

JETSET is written entirely in FORTRAN 77, and should run on any machine with a reasonable FORTRAN 77 compiler. Problems may exist on some machines. Thus the use of sequential access files for unformatted read/write is not allowed on the APOLLO, but this feature is anyway not strictly needed (it only appears in the optional storing of random number seed values).

More importantly, options for generating optimized code are known to cause problems on a number of machines, including IBM and DEC. These are *not* bugs in JETSET, but compiler bugs, and can be avoided by choosing a low optimization level. One typical problem, that appears on the IBM with optimization level 3, is that statements get pulled out of an IF ... THEN ... ELSE structure inside a DO loop, and placed before the loop, without the IF test being performed. Then an uninitialized variable may be used, leading to a program crash.

While version 6.3 required the user to interface and link an external random number generator, version 7.1 includes the new RANMAR [22] machine-independent, long-period random number generator, in a slightly modifed version compared to the CERN library one (a different seed value etc.). There is therefore no need to link to any external libraries.

All particle data and default values are stored in a BLOCK DATA subprogram, LUDATA, which appears at the end of the program file. This subprogram must be properly linked, something which may require extra machine-dependent commands.

Since most machines in current use are 32 bit ones, this is the precision normally assumed. A few pieces of code have therefore had to be written in double precision. On a 64 bit machine, it is not useful to have double precision at all; either this can be switched off by a compiler option or by minor changes in the code (as described in the manual).

A test program, LUTEST, is included in the JETSET package. It is disguised as a subroutine, so the user has to run a main program consisting of the two lines CALL LUTEST(0) and END to generate 600 events of varying types. If JETSET has not been properly installed, this program is likely to crash, or at least generate a number of erroneous events, which will then give error warnings.

JETSET is built as a slave system, run by subroutine calls from a user-written main program, with details set by commonblock switches and parameters. The complete list of switches/parameters is very long, essentially for three reasons:

- the physics model in itself is complicated and has many free parameters, in particular for the flavour description;
- the author is fairly conscientious in making even rather peripheral and seldom used parameters available in commonblocks, instead of just assigning values inline;
- JETSET contains a broad range of physics options, which e.g. means separate sets of parameters for string and independent fragmentation, or for matrix elements and parton showers.

Some of the switches and parameters most relevant for e^+e^- applications are collected in Table 10.

For each e^+e^- annihilation event to be generated, a call is to be made to the subroutine LUEEVT, which has two parameters. The first, KFL, is the requested flavour

Table 10: Main switches and parameters in JETSET 7.1. The description is rather simplified; more details and options are given in the program manual.

Parameter name	Default value	Meaning
MSTJ(104)	5	number of flavours allowed in $e^+e^- \to q\bar{q}$
MSTJ(107)	0	0/1 = off/on switch for initial state γ radiation
MSTJ(101)	5	QCD: 2 = second order matrix el., 5 = parton showers
MSTJ(42)	2	1/2 = conventional/coherent shower evolution
MSTJ(105)	1	0/1 = off/on switch for fragmentation and decays
MSTJ(1)	1	1/2 = string/independent fragmentation
MSTJ(11)	1	Lund fragmentation function, or other shapes
PARU(101)	0.0072974	α_{em}
PARU(102)	0.229	$\sin^2\theta_W$
PARJ(123)	92.4	m_Z used in electroweak cross-section (GeV)
PARJ(124)	2.55	Γ_Z used in electroweak cross-section (GeV)
PARJ(122)	0.5	$\Lambda_{\overline{MS}}$ for second order matrix el. and R_{QCD} (GeV)
PARJ(125)	0.02	y matrix element cut-off
PARJ(81)	0.40	Λ_{QCD} in shower evolution (GeV)
PARJ(82)	1.0	lower mass cut-off of shower evolution (GeV)
PARJ(1)	0.10	diquark/quark production ratio
PARJ(2)	0.30	s/u quark production ratio
PARJ(3)	0.40	extra strange diquark suppression
PARJ(4)	0.05	spin 1 diquark suppression factor
PARJ(11)	0.50	spin 1 fraction for light (u, d) mesons
PARJ(12)	0.60	spin 1 fraction for strange mesons
PARJ(13)	0.75	spin 1 fraction for charm and heavier mesons
PARJ(21)	0.35	width of quark transverse momentum dist. (GeV)
PARJ(41)	0.5	a parameter of Lund fragmentation function
PARJ(42)	0.9	b parameter of Lund fragmentation function (GeV^{-2})
PMAS(6,1)	60.	top quark mass (GeV)
PMAS(7,1)	120.	fourth generation low quark mass (GeV)
PMAS(8,1)	200.	fourth generation high quark mass (GeV)
PMAS(17,1)	60.	fourth generation χ lepton mass (GeV)

of the primary $q\bar{q}$ pair, with 1 - 6 corresponding to d, u, s, c, b and t, respectively, and 0 to the physically expected mixture. The second parameter, ECM, is the total CM energy of the event, given in GeV.

The typical generation time per 92 GeV e^+e^- event is 0.035 s (true) IBM 3090 time.

There are no particular problems associated with running JETSET. The modest

amount of initialization that may be needed is taken care of by the program automatically. Parameter values need not be fixed once and for all at the beginning of a run, but may be changed freely at any time. A user should note, however, that the order of indices in two-dimensional arrays is the opposite to what is used in other programs: in the momentum array, e.g., the first index is the particle number and the second the momentum component. Thus the components of a particle's four momentum are not stored in contiguous locations in computer memory, which can lead to inconvenience in the manipulation of the JETSET event record by the user.

Version 7.1 is based on the Particle Data Group standard particle numbering scheme, see section 6.1. A separate compressed code is used internally for the storage of particle properties, but is not visible to the outside observer.

4.8.7 Availability and Future Plans

JETSET version 6.3 is available in the CERN library, and can either be accessed in PATCHY format or as a ready-made TXTLIB; on the IBM do 'FIND JETSET' for further details. A non-PATCHY copy is stored on disk TORSJO 192 as JETSET63 FORTRAN. Since 6.3 still is the default version for many users, it is expected that version 6.3 will coexist for some time to come with later ones, i.e. 7.1 and beyond.

Version 7.1 is so far only available in JETSET71 FORTRAN on the TORSJO 192 disk, with the companion file JETSET71 MANUAL (4200 lines) containing the complete manual. The compiled version JETSET71 TEXT can be linked with an INCLUDE or a LOAD statement. The version on this disk may be gradually updated, if bugs are found, or for minor additions. When JETSET 7.2 appears, 7.1 will first be frozen, and later completely removed. Version 7.2 will be submitted to the CERN library.

From outside CERN, programs may obtained over BITNET by contacting the author on TORSJO@CERNVM.

Version 7.2 will probably appear during the course of 1989. Changes compared to 7.1 will be modest, and backwards compatibility will be kept, with some minor exceptions. A subroutine which performs conversion from the LUJETS commonblock to the proposed standard HEPEVT commonblock discussed in section 6.2, and the other way around, will be included (conceivably that routine could be added already to version 7.1). Further changes and additions may include

- photon emission as part of final state shower evolution;
- azimuthal asymmetries in shower evolution, as caused by soft gluon interference;
- more second order matrix element options, in addition to the GKS one;
- the possibility to pick an optimized scale for second order matrix elements; and
- a retuning of fragmentation parameters with tensor meson production included.

In the long term, it is planned to make use of the PYTHIA program [10] for the generation of the hard scattering process and to use JETSET only for the subsequent shower evolution, fragmentation, decays, and utility functions. Thus it will be possible to use the same matrix elements, e.g. to generate a W^+W^- pair, or to describe Higgs production via various processes, whether the incoming beams be e^+e^- or pp (or a few other alternatives). Much of the machinery for this already exists inside PYTHIA, although additional components will have to be added for a realistic description, in particular of threshold behaviour.

References

[1] T. Sjöstrand, Computer Physics Commun. **39** (1986) 347

[2] T. Sjöstrand, M. Bengtsson, Computer Physics Commun. **43** (1987) 367

[3] B. Andersson, G. Gustafson, C. Peterson, Z. Physik **C1** (1979) 105;
B. Andersson, G. Gustafson, Z. Physik **C3** (1980) 22;
B. Andersson, G. Gustafson, T. Sjöstrand, Z. Physik **C6** (1980) 235; Phys. Lett. **B94** (1980) 211; Nucl. Phys. **B197** (1982) 45; Z. Physik **C12** (1982) 49;
B. Andersson, G. Gustafson, B. Söderberg, Z. Physik **C20** (1983) 317

[4] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rep. **97** (1983) 31

[5] T. Sjöstrand, Phys. Lett. **B142** (1984) 420; Nucl. Phys **B248** (1984) 469

[6] B. Andersson, G. Gustafson, T. Sjöstrand, Physica Scripta **32** (1985) 574

[7] T. Sjöstrand, Z. Physik **C26** (1984) 93

[8] M. Bengtsson, T. Sjöstrand, Phys. Lett. **B185** (1987) 435; Nucl. Phys. **B289** (1987) 810

[9] F.A. Berends, R. Kleiss, Nucl. Phys. **B260** (1985) 32

[10] H.-U. Bengtsson, T. Sjöstrand, Computer Physics Commun. **46** (1987) 43

[11] G. Ingelman, T. Sjöstrand, LU TP 80-12 (1980);
G. Ingelman, LEPTO version 4.3, CERN program pool long writeup, program W5046 (1986)

[12] F.A. Berends, R. Kleiss, S. Jadach, Nucl. Phys. **B202** (1982) 63; Computer Physics Commun. **29** (1983) 185

[13] G. Bonneau, F. Martin, Nucl. Phys. **B27** (1971) 381

[14] B.R. Webber, Ann. Rev. Nucl. Part. Sci. **36** (1985) 253

[15] H.A. Olsen, P. Osland, I. Øverbø, Nucl. Phys. **B171** (1980) 209

[16] Particle Data Group, G. P. Yost *et. al.*, Phys. Lett. **B204** (1988) 1

[17] X. Artru, Phys. Rep. **97** (1983) 147

[18] B. Lörstad, Int. J. of Mod. Phys. **A4** (1989) 2861

[19] W.A. Zajc, Phys. Rev. **D35** (1987) 3396

[20] JADE Collaboration, W. Bartel *et. al.*, Z. Physik **C33** (1986) 23;
S. Bethke, Habilitation thesis, LBL 50-208 (1987)

[21] T. Sjöstrand, Int. J. of Mod. Phys. **A3** (1988) 751

[22] G. Marsaglia, A. Zaman, FSU-SCRI-87-50 (1987);
F. James, CERN-DD/88/22 (1988)

4.9 NLLjet

4.9.1 Basic Facts

Program name: NLLJET, a QCD parton shower Monte Carlo based on the Next to Leading Logarithmic Approximation [1]

Version: NLLJET 0.1, from March 1989

Authors: K. Kato
Physics Dept., Kougakuin Univ., Shinjuku-ku, Tokyo, Japan 160
T. Munehisa
Faculty of Engineering, Yamanashi Univ., Kofu, Yamanashi, Japan 400
T. Kamae
Dept. of Physics, Univ. of Tokyo, Bunkyo-ku, Tokyo, Japan 113
A.Shirahashi
Dept. of Physics, Univ. of Tokyo, Bunkyo-ku, Tokyo, Japan 113

Program size: 8266 lines

4.9.2 Physics Introduction

The Monte Carlo simulation of perturbative QCD is ordinarily based either on matrix elements or on the Leading Logarithmic Approximation (LLA), as discussed in section 2.2. The advantage of the former is the well defined nature of the theory, leading to quantitative predictions which may be tested experimentally. The disadvantage of matrix elements is that they exist to low orders only, due to the complexity of the calculations. Matrix element predictions exist only up to second order for the differential cross-sections which are necessary for a Monte Carlo simulation. Meanwhile various experimental studies have established the likely importance of multiple QCD radiation for the description of the overall characteristics of multihadronic annihilation data at LEP energies [2]. In contrast, Monte Carlo simulations based on the Leading Logarithmic Approximation *do* possess multiple parton emission terms. However QCD to the leading logarithmic level contains many ambiguities not present in the matrix element formulations. These ambiguities include the functional forms of the arguments Q and z of the strong coupling constant $\alpha_S(Q^2)$ and of the Altarelli-Parisi splitting kernels $P_{i\to jk}(z)$. In addition, the renormalization scheme is left ambiguous in LLA models, in that calculations employing different renormalization schemes are equivalent to the leading logarithmic level. As a consequence, quantitative tests of QCD are difficult with leading logarithm based Monte Carlos. In particular, the QCD scale variable Λ, which is present as a fundamental parameter in the LLA models, cannot be calculated theoretically. Therefore the values of Λ extracted from LLA models through comparision with data have no real theoretical meaning.

The NLLJET Monte Carlo of Kato *et. al.* attempts to remedy this situation through the construction of a parton shower algorithm based on the Next to Leading Logarithmic Approximation (NLLA), which includes one-loop corrections to the LLA picture.

At the NLLA level, the theoretical ambiguities discussed in the previous paragraph are no longer present. The renormalization scheme may be specified unambiguously, for example. As a consequence, NLLJET holds the promise for a QCD Monte Carlo algorithm which possesses the advantages both of matrix elements (well defined theoretical input permitting a quantitative test of QCD) and of parton showers (incorporation of multiple parton emission and a better overall description of data at high energies).

It should be stated that NLLJET is in an active state of development. The authors do not claim that the full promise of their program has yet been realized. The computer code for NLLJET has not been 'officially' released to the general user. One aspect of the program which the authors feel requires further work is the inclusion of the $q\bar{q}g$ state from the electroweak vertex (in e^+e^- annihilations), to supplement the $q\bar{q}$ state, in a manner which is theoretically consistent and which has been adequately tested. Such work will certainly be necessary before the claim may be substantiated that the 'Λ parameter' of NLLJET is equivalent to $\Lambda_{\overline{MS}}$.

4.9.3 Detailed Physics Description

Electroweak cross-section and QED radiative corrections:
The electroweak point-like scattering in NLLJET is based on first order electroweak matrix elements describing the exchange of a photon or a Z^0 in the s channel. At most one initial-state photon may thereby be emitted in the scattering process. Quark mass effects are included in these expresssions. The electroweak cross-section and QED radiative corrections implemented in NLLJET are therefore similar to those of JETSET. The authors of NLLJET state that they intend to extend their program to include higher order QED radiation in the electroweak scattering at some point in the future.

It is possible with NLLJET to specify the creation of a specific quark flavour or to generate a mixture of flavours according to the relative values of the cross-sections for the individual quarks.

The Next to Leading Logarithmic Parton Shower:
The physics behind the Next to Leading Logarithmic Approximation as implemented into the NLLJET parton shower algorithm is presented in refs. [3,4], which also contains references to the underlying theoretical work. Briefly, the Next to Leading Logarithmic Approximation is a summation over terms which represent the mass singularities in a perturbative expansion – as is the usual LLA – but it improves on the LLA by including parton branchings which are valid to the second order rather than to just the first order of the strong coupling constant. The NLLA thus includes the three body parton splittings

$$\begin{aligned} q &\to q+g+g, \\ q &\to q+q'+\bar{q}', \\ g &\to g+g+g, \\ g &\to g+q+\bar{q}, \end{aligned} \tag{148}$$

where q represents a quark and g a gluon. The two body parton splittings,

$$\begin{aligned} q &\to q+g, \\ g &\to g+g, \end{aligned}$$

$$g \rightarrow q + \overline{q}, \tag{149}$$

are of course also included in the NLLA summation, with the Altarelli-Parisi splitting functions $P_{a\to bc}$ here given to second order in α_S (i.e. including one-loop effects). In contrast, the usual Leading Logarithmic Approximation includes only the parton splittings (149), with splitting functions only given to first order in α_S. The Next to Leading Logarithmic Approximation might thus be said to hold the same position relative to the Leading Logarithmic Approximation that second order matrix elements hold to the first order expressions.

The Monte Carlo algorithm itself is discussed somewhat in [1]. In its general principles this algorithm is not so different from those of other QCD cascade Monte Carlos (see the description of the parton shower simulation in HERWIG, for example); however there is necessarily much richness of detail in the implementation because of the multitude of branching possibilities (148) and (149).

The main evolution variable is the parton virtuality, i.e. $Q^2_{evol} = m^2$. At an $a \rightarrow bc$ branching, the z variable is chosen as the light-cone fraction

$$z = \frac{E_b + p_{Lb}}{E_a + p_{La}}, \tag{150}$$

where the longitudinal momentum direction is chosen along the a parton momentum vector. At an $a \rightarrow bcd$ branching, two variables z_b and z_c are to be selected, with $z_d = 1 - z_b - z_c$ by energy and momentum conservation. In addition, the full kinematics is rather more complicated.

In its earliest versions, the NLLJET showering algorithm used a gauge choice such that only one of the initial partons (q or $\overline{q}$) was free to radiate [3]. In the current version, both initial partons are on equal footing and may radiate [4]. Due to the light-cone z definition, a low z value corresponds to a parton emitted in the backwards hemisphere. To avoid doublecounting in the total emission probability, it is therefore necessary to impose an extra cut on allowed z values, a cut which depends on the virtuality of the emitting parton and which uniquely separates q from $\overline{q}$ emission.

The NLLJET Monte Carlo includes angular ordering in the polar emission angles of the partons in its shower, in order to simulate some of the consequences of interference due to the coherent emission of soft gluons.

Fragmentation and Decays:
Originally, a separate fragmentation package was used. In this model, gluons were split into $q'\overline{q}'$ pairs, with final quarks rearranged into colour singlet systems (just like e.g. in HERWIG). Each $q_1\overline{q}_2$ system was then fragmented like a pair of back-to-back jets, in the spirit of independent/string fragmentation models (the two essentially agree for this simple system), with special provisos for the handling of low-mass systems.

In the current version, the fragmentation of partons and the decays of hadrons are performed using the JETSET library, version 6.3.

4.9.4 Comparison with Data

The authors of NLLJET have compared the predictions of their Monte Carlo, for global multihadronic event shapes (thrust, sphericity etc.), for the energy-energy correlation

and its asymmetry and for jet multiplicites, to experimental data in e^+e^- annihilations at CM energies between 29 GeV and 60 GeV [1,5]. Generally they find quite good agreement between model and data except for the energy-energy correlation asymmetry (EECA). They suggest that the problem with the EECA is a consequence of the lack of a complete integration of the matrix element for $e^+e^- \rightarrow q\overline{q}g$ into their model [1].

4.9.5 Installation and Availability

The NLLJET program has not yet been released to the general public. The authors first wish to further develop and test their program. BITNET mail may be addressed to KAMAE@JPNKEKVM or DECNET mail to KEKVAX::KAMAE or to TKYVAX::KAMAE for inquiries, however.

The current version of the program (version 0.1) is written in VAX specific Fortran but is convertable to run on IBM VM/XA with relatively modest effort. The program initialization time is very long, about twenty IBM 3090 minutes or about two VAX 8530 hours. This time is (mainly) spent in tabulating the Sudakov form factors needed in the shower evolution. The authors intend to reduce this time in the future. A feature exists by which the start-up data may be stored on disk so as to avoid the long initialization time, except for the first time the program is run. Even when making use of this feature the initialization time is several IBM 3090 minutes, however. The long initialization time cannot be avoided should either one of the main perturbative parameters controlling the parton shower or the CM energy be changed.

References

[1] 'Next-to-LL Parton Shower Generator for e^+e^- Hadronic Annihilation – Description of the program NLLjet', K. Kato, T. Munehisa, T. Kamae, A. Shirahashi, March 1989 (unpublished);
T. Kamae *et. al.*, 'Study of Hadron Jets in e^+e^- by a New Simulation Program Including the Next-to-LL Order', Contribution 0779B to the 24th Int. Conf. on High Energy Physics, Munich 1988 (unpublished)

[2] JADE Collaboration, W. Bartel *et. al.*, Z. Physik **C33** (1986) 23;
TASSO Collaboration, W. Braunschweig *et. al.*, Phys. Lett. **B214** (1988) 286;
TPC/2γ Collaboration, H. Aihara *et. al.*, Univ. of Tokyo preprint UT-HE-89/02

[3] K. Kato, T. Munehisa, Mod. Phys. Lett. **A1** (1986) 345, Phys. Rev. **D36** (1987) 61

[4] K. Kato, T. Munehisa, Phys. Rev. **D39** (1989) 156

[5] T. Kamae, Proceedings of the 24th Int. Conf. on High Energy Physics, Munich 1988.

4.10 PARJET

4.10.1 Basic Facts

Program name: PARJET for parton shower evolution [1,2,3]
BAMJET for fragmentation [4,5]
DECAY for decays [6]
Version: PARJET-88 from 1988 [7]
Authors: Fippel, K. Haensgen, R. Kirschner, J. Ranft, S. Ritter
Arbeitsgruppe Hochenergiephysik
Sektion Physik
Karl Marx Universität
Karl Marx-Platz 10-11
DDR-701 Leipzig
BITNET RAN @ CERNVM (J. Ranft; only occasionally at CERN)
Program size: 4542 lines

4.10.2 Physics Introduction

The PARJET code of Ritter [1,2,3] implements the evolution of quark-antiquark jets, as produced in e^+e^-. The parton decay into partons is treated on the basis of perturbative QCD in the leading logarithmic approximation (LLA), according to the procedures developed in Refs. [2,3,8,9]. The LLA allows a simple probabilistic interpretation of the parton evolution as a branching process. The branching process is stopped if the invariant masses of the decaying partons become less than a non-perturbative cut-off parameter μ, in the order of a typical hadronic mass scale. PARJET was modified in 1988 by Kirschner and Fippel [7], so as to include the angular ordering scheme [10,11,12].

PARJET uses for the fragmentation into hadrons the code BAMJET [4,5], and for the decay of hadron resonances the code DECAY [6]. Both are included in the PARJET package. The Refs. [1,4,6] contain the long writeups of these procedures.

4.10.3 Comparison with Data

A comparison of the results obtained with the Monte Carlo code PARJET (without angular ordering) has been published in Ref. [3] (1982). Good agreement with data from e^+e^- annihilation into hadrons from the PETRA storage ring in the energy range up to $\sqrt{s} \approx 30$ GeV was found. Studied were in particular average charged multiplicities, charged multiplicity distributions, average multiplicity for charged pions, charged kaons and antiprotons, average transverse and longitudinal momenta, and average transverse momentum-squared of charged particles, all as a function of $\sqrt{s}$; further transverse momentum distributions dN/dp_T^2 of all charged particles, the cross-section $s\,d\sigma/dz$ as a function of scaled hadron momentum fraction $z = 2p/\sqrt{s}$ for charged particles, rapidity distributions, inclusive cross-sections for $\pi^\pm$, $K^\pm$, $p\overline{p}$ and $\Lambda\overline{\Lambda}$ production, energy-energy correlations, and charge compensation probabilities as a function of rapidity.

4.10.4 The code: Availability and Plans

The Monte Carlo code PARJET, originally written in FORTRAN 4, runs in its present version under FORTRAN 77. It has been tested on the IBM mainframe computer. For 1000 events at $\sqrt{s} \approx 100$ GeV, one needs about 1 minute of CPU time on the IBM 3090.

The most important input parameters are:

- centre of mass energy $\sqrt{s}$ of the $q\overline{q}$ system;
- initial quark flavour (in the present version only $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, and $c\overline{c}$ initial states are possible);
- QCD scale parameter Λ;
- minimal allowed final parton mass μ.

The fragmentation code BAMJET is controlled by some parameters, which allow a modification of the fragmentation, see Ref. [4]. The published writeups [1,4,6] are valid also for the new code PARJET-88.

Present plans include a detailed comparison of PARJET-88 with data and a modernizing of the fragmentation and decay codes, especially regarding charm and bottom decays.

The code is freely available, and can be found on disk RAN 195 on the CERNVM system, as follows:

- the file TEBAMPAR FORTRAN contains the program proper;
- the file TEBAMPAR DATA contains necessary runtime data; and
- the file BAMPAR EXEC contains a sample command file for running the program.

References

[1] S. Ritter, Computer Physics Commun. **31** (1984) 401
[2] R. Kirschner, S. Ritter, Physica Scripta **23** (1981) 763
[3] S. Ritter, Z. Physik **C16** (1982) 27
[4] S. Ritter, Computer Physics Commun. **31** (1984) 393
[5] J. Ranft, S. Ritter, Acta Phys. Pol. **B11** (1980) 259
[6] K. Haensgen, S. Ritter, Computer Physics Commun. **31** (1984) 411
[7] Fippel, R. Kirschner, in preparation
[8] G.C. Fox, S. Wolfram, Nucl. Phys. **B168** (1980) 285, Z. Physik **C4** (1980) 237
[9] R. Odorico, Nucl.Phys. **B172** (1980) 157
[10] A.H. Mueller, Phys. Lett. **B104** (1981) 161
[11] A. Bassetto, M. Ciafaloni, G. Marchesini, A.H. Mueller, Nucl. Phys. **B207** (1982) 189
[12] Yu.L. Dokshitzer, V.S. Fadin, V.A. Khoze, Phys. Lett. **B115** (1982) 242

4.11 TIPTOP

4.11.1 Basic Facts

Program name: TIPTOP, a Monte Carlo for Heavy Fermion Production and Decay in e^+e^- Annihilation [1,2];
TIPLOT, a Monte Carlo for Forward-Backward Asymmetries in Heavy Fermion Production

Version: TIPTOP version 1.1, TIPLOT version 1.1

Authors: S. Jadach
Institute of Physics
Jagellonian University
Cracow, Poland
BITNET JADACH @ CERNVM (when at CERN)
Johann H. Kühn
Max Planck Institute for Physics
München, Fed. Rep. Germany
BITNET JXK @ DM0MPI11

Program size: TIPTOP 2150 lines, TIPLOT 387 lines

4.11.2 Physics Introduction

One of the important aims of LEP will be the search for heavy fermions. Besides the top quark, these heavy fermions include a fourth generation down type quark b', a fourth generation lepton L, as well as other coloured fermions predicted by more exotic models. It is thus important to have a Monte Carlo that gives a precise simulation of the production and the decay of these fermions. Because of the short lifetime of heavy fermions, it is convenient to treat production and decay jointly.

A description should take into account the electrodynamic radiative corrections, which are known to give important contributions to lepton and quark production. Ideally, the Monte Carlo should also include QCD corrections. Finally, the semileptonic and hadronic decay modes require the use of a hadronization scheme.

The Monte Carlo TIPTOP is designed to fulfill some of the requirements above by having the following features:

1. It simulates the production process $e^+e^- \to F\overline{F}(\gamma)$ followed by the decay $F \to f_1 + f_2 + \overline{f}_3$.
2. It incorporates initial state bremsstrahlung.
3. It takes into account spin and mass effects of F and the weak isopartner f_1.
4. It has a switch for simulation of the production and decay of a heavy lepton L, top quark t, or fourth generation b', followed by their decays.
5. It takes into account polarization of F, and its effect on the decay products.
6. It takes into account longitudinal beam polarization.
7. For $m_F < m_W$ the effect of the W propagator are included, and for $m_F > m_W$ the production of a real W with finite width, with subsequent decay including spin effects.

The program is a natural extension of KORALB of Jadach and Wąs [3,4] for τ production, generalized to include heavy quarks. It has thus the important features of the inclusion of mass and spin effects of that program. The TIPTOP package also comes with a program to calculate the forward-backward asymmetries by taking into account initial state radiation, quark polarization, and kinematical effects on the heavy quark decay. In particular, spin effects contribute to these asymmetries.

The program package consists of the five distinct and relatively well separated parts:

1. EMAUS Monte Carlo for the production process $e^+e^- \to F\overline{F}(\gamma)$.
2. TOPOLA for the decay $F \to f_1 + f_2 + \overline{f}_3$.
3. MATRIX ELEMENT for EMAUS.
4. JMCLIB general purpose library of Monte Carlo and kinematics.
5. TIPLOT Monte Carlo for calculating forward-backward asymmetries.

The detailed description of each of these programs follows.

4.11.3 Detailed Physics Description

Production of the fermions:
The heavy fermion production process $e^+e^- \to F\overline{F}(\gamma)$ is very complicated from the simulation point of view [5]. This is because the differential cross-sections does not factorize into production and decay parts, heavy fermion decay being sensitive to spin polarization.This difficulty is resolved in TIPTOP in the following fashion.

First, the $e^+e^- \to F\overline{F}$ event is generated, assuming that the incoming electron/positron pair is unpolarized, and similarly the fermion decays are simulated with the neglect of spin polarization effects. Only afterwards is any e^+e^- polarization dependence introduced, together with all polarization effects in the fermion decays, in the form of one singe overall event weight, which is used to decide whether to keep or reject the event.

This part of the program is based on the program KORALB for the production of τ pairs. The amplitude for the process $e^+e^- \to F\overline{F}$ is written in a similar fashion as that for τ pairs [2], i.e.:

$$\begin{aligned} T &= \frac{e^2}{s}\Big[\overline{u}(F)Q_F\gamma_\mu v(\overline{F})\overline{v}(e^+)Q_e\gamma^\mu u(e^-)+ \\ &\quad \chi(s)\overline{u}(F)(v_F + a_F\gamma_5)\gamma_\mu v(\overline{F})\overline{v}(e^+)(v_e + a_e\gamma_5)e\gamma^\mu u(e^-)\Big], \end{aligned} \tag{151}$$

where $\chi(s)$ is the ratio of Z^0 and γ propagators, eq. (5).

The novel feature of the program is the fact that spin effects are included. The cross-section is given by:

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_0 (1 + a_\mu s^\mu + \overline{a}_\mu \overline{s}^\mu + c_{\mu\nu}s^\mu s^\nu). \tag{152}$$

Here $(d\sigma/d\Omega)_0$ stands for the spin averaged differential cross-section, and s^μ and $\overline{s}^\mu$ for the spin vectors of F and $\overline{F}$, respectively. The a_μ, $\overline{a}_\mu$ and $c_{\mu\nu}$ are functions of the four-momenta of the initial and final particles, including any photon. The explicit formulae are given in ref [6].

Initial state radiation is also included in the production part of the Monte Carlo, EMAUS. If a photon is emitted, the four-momenta of the F and $\overline{F}$ are chosen in the reduced center of mass system, according to the lowest order $d\sigma/d\Omega$.

Initial state QED radiation:
In the first part of the production algorithm, the photon momentum E_γ is generated according to the distribution:

$$\frac{d\sigma}{dk} = \sigma_0 \left[\delta(k)\rho_s(k_0) + \Theta(k-k_0)\rho(k)\right], \tag{153}$$

where $k = E_\gamma/E_{beam}$. The pointlike cross-section $\sigma_0 = 4\pi\alpha^2 e^4/3s$ is used as normalization. A cut at $k_0 = 0.003$ (called EPS in the program) is introduced in the photon energy spectrum, so as to isolate the infrared singularity. Hard events, with $k > k_0$, and soft events are handled separately. The function ρ_s collects contributions from the lowest order virtual and soft radiation graphs, whereas $\rho(k)$ gives the hard contributions.

The method used for the generation of the photon is the weighting rejection technique described in [5].

Inclusion of spin effects:
The production cross-section, including the dependence of the spin vectors, is given in eq. 152. The decay amplitude(s), given in the appendix of [1], can be cast in the form

$$\begin{aligned} d\Gamma_F &= (1 + b_\mu s^\mu)(d\Gamma_F)_0, \\ d\Gamma_{\overline{F}} &= (1 + \overline{b}_\mu \overline{s}^\mu)(d\Gamma_{\overline{F}})_0. \end{aligned} \tag{154}$$

Here $(d\Gamma)_0$ is the spin independent cross-section, which is a function of the initial and final momenta.

The second part of the program, TOPOLA, first simulates the decays as the F and the $\overline{F}$ unpolarized. It also calculates the vectors b and $\overline{b}$, and transfers them back to EMAUS. At this stage, the $e^\pm$ polarization effects and the spin effects in the fermion decay are introduced as a single weight, which is used to reject events. The resulting multi-differential cross-section for the combined production and decay process is given by

$$d\sigma_{spin} = \left(1 - a_\mu b^\mu - \overline{a}_\mu \overline{b}^\mu + c_{\mu\nu} b^\mu \overline{b}^\nu\right)(d\sigma)_0 (d\Gamma_F)_0 (d\Gamma_{\overline{F}})_0 \tag{155}$$

The complete description of this technique is given in [3].

Polarization effects may enhance or deplete distributions of decay products of one fermion by a factor between one and two, and of two fermions between one and four. For heavy quarks, it must be noted that hadronization may alter the spin dependence of the production and decay cross-sections. In general, one assumes that the form of these terms is not changed, but that only their magnitude is reduced by a factor f [7]:

$$d\sigma_{final} = \left[1 - f\left(a_\mu b^\mu + \overline{a}_\mu \overline{b}^\mu\right) + f^2 c_{\mu\nu} b^\mu \overline{b}^\nu\right](d\sigma)_0 (d\Gamma_F)_0 (d\Gamma_{\overline{F}})_0 D_F D_{\overline{F}}, \tag{156}$$

where D_F and $D_{\overline{F}}$ are the fragmentation functions for F and $\overline{F}$, see section 2.3 and the JETSET description.

Fragmentation:
Fragmentation is performed by calls to the JETSET program.

Forward-backward asymmetries:
Forward-backward asymmetries constitute important measurements at LEP, since they serve as tests for the electroweak couplings of heavy quarks, and of any new couplings

envisaged beyond the standard model. Two types of asymmetries can be studied, using TIPLOT in conjunction with TIPTOP: either the asymmetries of the hevy quarks themselves, by reconstruction from their decay products, or, more directly, the asymmetries of the leptonic final states. Experimentally, the first method is subject to the uncertainty arising from the choice of reconstruction technique, as well as the effects of QCD corrections, but it is directly related to the asymmetry. The second method is cleaner, but is less direct.

The study of forward-backward asymmetries of decay product leptons avoids the complications of QCD corrections. However, because of the energetics of the decay process, there is a spillover of decay products into the opposite hemisphere, which affects the lepton asymmetries. The spin polarization of the quarks also affects results. All this can be well studied by this Monte Carlo. Initial state radiation effects can also be studied, and are seen to affect the quark and lepton asymmetries only slightly. As mentioned above, in the hadronization process the spin polarization is assumed to be diluted by a factor f. The dependence of the asymmetries on f can and has been studied by this Monte Carlo; thus various models can be tested by studying both the angular and energy distributions of heavy quark decays.

4.11.4 Limitations of the Program

The major limitations of TIPTOP are the following:

1. It does not include QCD corrections, which may have important effects on the shape of the lepton spectrum and on the total decay rate of heavy quarks. In particular, gluon radiation effects are not included.
2. Final state QED radiation is not included.
3. The $f_1 f_2$ isodoublets couple in a pure $V - A$ fashion.

4.11.5 Comparison with Data

Although, of course, no data is presently available to compare the results of this Monte Carlo for top and b' production, several test runs have been made by the authors to test the effects of initial state radiation and fermion polarization on the single muon energy spectrum of the decay of the heavy lepton/quark. Muons are studied so as to avoid the complications mentioned earlier about the choice of the fragmentation function. For a lepton mass of the order of 75% of beam energy E (tested at $E = 80$ GeV) the initial state radiation has negligible consequences. However, as the lepton mass becomes a smaller fraction of the beam energy, the effects are more pronounced.

Fermion polarization, on the other hand, has more dramatic effects on the single muon distribution. This is because in Z^0 decay into heavy fermions, the polarization is large and the distributions reflect this. The authors have studied the effect for $m_t = 40$ GeV and three values for the depolarisation factor, $f = 0, 1/2, 1$ (see ref. [1]).

The same studies have been performed for the forward-backward asymmetries, both for the heavy fermions and for the prompt muon decays. The effect of the initial state radiation is much more pronounced in this case, see ref [2].

4.11.6 Installation and Running

The following programs are installed on the CERNVM, on disk BAMBAH 191. They can also be obtained directly from the authour Kühn.

1. TIPDEMO FORTRAN: The main demonstration program.
2. TIPTOP FORTRAN : The TIPTOP program itself.
3. TIPDEMO DATA : Input data.
4. TIPDEMO LISTING : Output listing from TIPDEMO.

The TIPDEMO is an example of how to use the program as a black box. By default, it generates heavy lepton events.

A complete e^+e^- event is generated by calls to the subroutine TIPTOP(MODE,IDPROC,ENE,IPAR,PAR), where the first parameter MODE is -1, 0 or +1 depending on whether TIPTOP is in the initialization mode, the generation mode, or the termination mode. The Parameter IDPROC selects the heavy fermion type. IDPROC = 2, 3 and 4 corresponds to a heavy lepton, top and b', respectively. ENE is the beam energy. The array IPAR(1-6) contains switches which can turn initial state radiation and spin effects on and off. The array PAR(1-16) contains information about polarization, fragmentation depolarization factor, particle masses, and various couplings. The assignments are well explained in the program.

In the generation/production mode, MODE = 0 , the events produced are stored in the common LUJETS. The hadronization takes place by a call to LUEXEC. The demonstration program TIPDEMO provided with TIPTOP can be run without LUJETS by deactivating the commented lines. It should be noted that the original program uses the old version of HBOOK – this has been altered in the files listed above.

In the post-generation mode, MODE = 1, information concerning the integrated cross-section and the partial decay rates of the heavy fermions can be printed in the output files and can be read through the arrays PAR and IPAR.

References

[1] S. Jadach, J. H. Kühn, 'TIPTOP, a Monte-Carlo for Heavy Fermion Production and Decay', MPI PAE Pth 86-64 (1986)

[2] S. Jadach, J.H. Kühn, Phys. Lett. **B191** (1987) 313

[3] S. Jadach, Z. Wąs, Computer Physics Commun. **36** (1986) 191

[4] S. Jadach, Acta Physica Polonica **B16** (1985) 1007

[5] F. A. Berends, R. Kleiss, S. Jadach, Z. Wąs, Acta Physica Polonica **B14** (1983) 413

[6] J. H. Kühn, Nucl. Phys. **B272** (1986) 560

[7] J. H. Kühn, Nucl. Phys. **B237** (1984) 77

5 Physics Issues

This section has three main purposes:

1. to give an overview of experience gained in comparing e^+e^- data with existing event generator models;
2. to remind the reader of unresolved issues, that maybe have not been considered in programs development, or are present only at a very primitive level; and
3. to point at experimental studies at LEP that could help us increase our understanding of multihadronic events and the modelling necessary to describe them.

In the first half of the section, the emphasis is mainly on the experience gained in the past, while the latter half is more directed towards the future and unresolved issues.

5.1 On Comparisons between Models and Data

Hadron distributions in a jet are due to soft interactions and cannot be calculated from first principles. Instead one has to rely on models which are more or less inspired by QCD. These models provide an essential tool for understanding how hadronic events evolve. The exploration of the structure of jets was one of the main activities at high energy e^+e^- colliders in recent years. The sizable event rates at PETRA and PEP at Q^2's around 1000 GeV^2 boosted our insight into jet properties. In this section we review the published comparisons between fragmentation models and the data collected at PETRA, PEP and TRISTAN. More details on QCD analyses in e^+e^- experiments and on the measurements of the structure of jets in e^+e^- collisions can be found in [1].

Comparing models and data is complicated due to the existence of more or less many free parameters. These are frequently not related to the basic ideas of the model, but affect the final distributions strongly. The number of free parameters differs quite a lot among the various models. In most cases these parametrizations cannot be measured directly but have to be inferred from the data in an indirect way. Flaws in the basic concepts of a model can at least in part be covered by a favourable parametrization. Also the reverse is true: an unappropriate choice of fragmentation parameters can spoil the evaluation of a model: a basically correct concept can be put aside because of that. To evaluate the validity of a certain concept requires therefore to find distributions which reflect it most directly and to separate it from the uncertainties in the details of hadronization.

The importance of knowing the structure of hadron production goes beyond the search for a correct description of QCD at low Q^2. Being able to predict and estimate the features of hadronic events is important for extracting information about the production of new particles, for tagging the primary parton type and in such a way measuring quark couplings, or for unravelling fundamentals of high Q^2 QCD properties. Apart from trying out basic concepts for hard and soft hadronization effects, these models may serve the additional purpose of just having a parametrizations of the data.

The comparisons of data and models can be crudely subdivided into those testing the underlying more or less hard parton structure and those testing its transformation into hadrons. The former is predominantly done using global event features, the latter using details of the particle flow and particle content.

In the following we will concentrate on the most frequently used models. Alternative

appoaches not based on QCD ideas exist and they sometimes reproduce specific measurements. In most cases they are not cast into a detailed simulation code and they are therefore difficult to assess. Obviously a good candidate for 'the' fragmentation model has to describe all features of the data and not only a few.

5.2 General Features of the Event Topology at $W = 30$ GeV

Data and model predictions have been compared in a much detail for distributions characterizing the overall event topology. To evaluate the validity of the various models the uncertainties from the free parameters were minimized. They were determined by a multi-parameter fit to sensitive distributions. On the basis of these optimized parameters the underlying concepts can be meaningfully compared.

Such analyses have been presented by the Mark II [2] and TASSO Collaborations [3]. The values obtained for the various models from this optimization procedure at CM energies W of 29 and 35 GeV, respectively, are listed in Tables 11 - 14. In general the fitted values agree quite well for the two experiments with the exception of the parametrizations of the symmetric Lund fragmentation function. One should also note that the number of free parameters is larger for the JETSET (Lund) [4] and the CALTECH [5] model than for Webber (i.e. the EARWIG predecessor to HERWIG) [6]. This reflects the two basically different ways of converting the generated parton distribution into the final hadrons in these models. The latter applies the rather simple assumption of a phase-space decay of clusters. The final distribution is determined essentially by the cluster mass. In the Lund model the hadronization occurs by break-ups of the string requiring several parameters for the fragmentation functions in addition to those determining the parton distributions. The CALTECH model uses a mixture of those two.

One should note the very low cut-off masses for the parton evolution in the Webber and JETSET shower models. They lead to a maximum number of individual generated partons before hadronization sets in.

On the l[illegible] of these parametrizations Mark II and TASSO analyzed the overall features of events. To cover the different facettes of fragmentation properties, distributions reflecting specific aspects were selected (some of them are displayed in Fig. 28): the broadness of events is measured by the thrust and sphericity being sensitive to hard QCD bremsstrahlung, the aplanarity is sensitive to soft effects and higher order hard QCD processes, the particle multiplicity and fragmentation function $f(x)$ with $x = 2p/W$, the scaled momentum, reflects the impact of soft processes.

The general conclusions from these comparisons are

- The CALTECH-II model (only considered by the Mark II Collaboration) describes the data poorly. The predictions for the distributions related to the transverse momentum p_T both in and out of the event plane have deficiencies: there are too many high aplanarity events, too many events of very high thrust and too many at very low thrust.
- The Webber Monte Carlo describes quantities well which are related to the p_T out of the event plane, whereas distributions related to p_T in the event plane are predicted to be too broad.
- The JETSET option based on the second order QCD matrix element gives a good description for event and particle properties in the plane, but has a deficiency of events of high aplanarity.
- The JETSET option using the first order QCD matrix element supplemented by a shower approach for higher parton multiplicities gives the best description of the data. Both the components in and out of the event plane are well described.

The results of the fits are expressed in terms of χ^2. Using 450 data points, the Mark II Collaboration finds a χ^2 of 960 for the JETSET shower option, and 1230 for the JETSET $\mathcal{O}(\alpha_S^2)$ option. The Webber simulation gives 2870 and CALTECH-II gives the worst χ^2 of 6830. Similar results were obtained by TASSO. These numbers should only be seen as indicators of some trend but not as exact statistical probabilities. They depend on what distributions have been used for their determination. In addition these

Table 11: Optimized parameters for the Webber model.

type	Mark II	TASSO
Λ_{LLA} (GeV)	0.2	0.25
m_g (GeV) cut-off mass	0.75	0.61
M_c (GeV) Cluster mass	3.0	2.3

Table 12: Optimized parameters for the CALTECH-II model.

type	*Mark II*	*TASSO*
Λ_{LLA} (GeV)	0.5	
$t_0(GeV^2)$ cut-off	2.0	
$\rho(GeV^{-2})$ string breaking	1.6	
$w_{max}(GeV)$ cluster decay	2.2	
$w_{min}(GeV)$ cluster decay	0.25	

Table 13: Optimized parameters for the JETSET shower option.

type	*Mark II*	*TASSO*
Λ_{LLA} (GeV)	0.4	0.26
$Q_0(GeV)$ cut-off	1.0	1.0
a(fragmentation function)	0.45	0.18
b(fragmentation function)	0.90	0.34
$\sigma_q(GeV)$ for p_T	0.33	0.28

Table 14: Optimized parameters for the JETSET $\mathcal{O}(\alpha_S^2)$ option.

type	*Mark II*	*TASSO*
$\Lambda_{\overline{MS}}$ (GeV)	0.5	0.62
y_{min} cut-off	0.015	0.02
a(fragmentation function)	0.9	0.58
b(fragmentation function)	0.7	0.41
$\sigma_q(GeV)$ for p_T	0.33	0.28

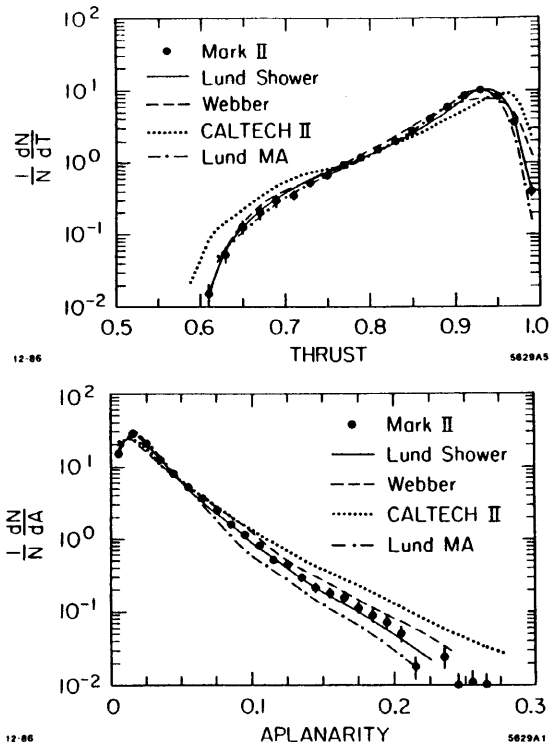

Figure 28: Thrust and aplanarity distributions, Mark II data [2] compared with model results.

are strongly correlated, something which is not properly taken into account in the χ^2 calculation.

A similar assessment of the various models has been obtained by studying the frequency of cluster multiplicities. In a first measurement [7] the observed n-cluster rates were compared to the JETSET ($\mathcal{O}(\alpha_S^2)$) option and to the shower model of Webber. It was found that this JETSET option predicts too few four and five cluster events at W = 34 GeV. On the other hand the shower model of Webber does reproduce these rates quite well, but underestimates those for the three-cluster rate and overestimates those for two clusters. These trends were confirmed in reference [8]. In addition it was shown there that the JETSET option based on a mixture of first order QCD matrix element calculation and subsequent showering leads to a much better agreement with the data for all cluster multiplicities (Fig. 29).

In summary these results suggest the following trend.

- The exact matrix element calculations up to the second order as implemented in the JETSET program underestimates the number of partons. This can be inferred from the deficiency of events with a large component out of the event plane or more directly from the cluster multiplicity. Higher order QCD corrections seem to be necessary. Their calculation is a formidable task and has as yet only been achieved either on the tree level or for the inclusive total hadronic cross-section.

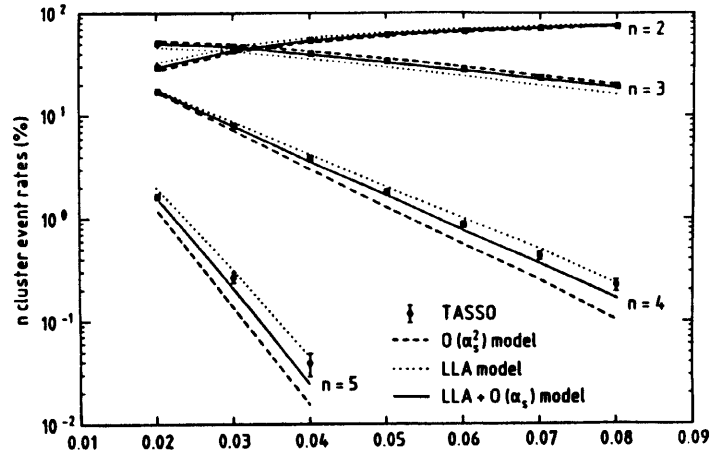

Figure 29: The n-jet event rates (%), defined using a clustering algorithm, as a function of the y resolution parameter of the algorithm; TASSO data [8] compared with model results (see text for explanations).

Recent results weaken this conclusion: it has been shown [9] that second order QCD corrections describe the data well if the renormalization point is optimized.

- The other approach is to calculate higher orders using the Leading Log formalism. This is the basis for the shower approaches. They give a good account of high jet multiplicities.

Taken at face value, the results discussed above suggest some caution: the deficiencies of the pure leading log approach in the model of Webber could be cured by forcing the first branching to occur according to the exact matrix element. One may infer that the shower development alone does not exactly reproduce the true parton distribution. However, the differences may also be due to the different fragmentation schemes applied in the two models.

Although the JETSET option combining first order QCD matrix elements and Leading Log Approximation to generate additional partons seems to be the best candidate for the structure of hadronic events at LEP energies, the recent results using the second order matrix element at a different Q^2 scale indicate a viable alternative. It will be interesting to see how the multiplicities and the differential distributions of partons at the Z^0 will agree with the results of the various approaches to multi-parton production.

5.3 The Q^2 Dependence of Inclusive Distributions

Apart from detailed studies at a fixed energy the energy dependence of fragmentation properties has been analysed. Data from PETRA exist for various CM energies between W = 14 GeV and 46 GeV. Recent results from TRISTAN extend the energy range of data on jet properties up to W = 56 GeV. This large span of CM energies allows one to check also the Q^2 dependence of the model predictions. This is especially important for extrapolating jet properties to LEP energies.

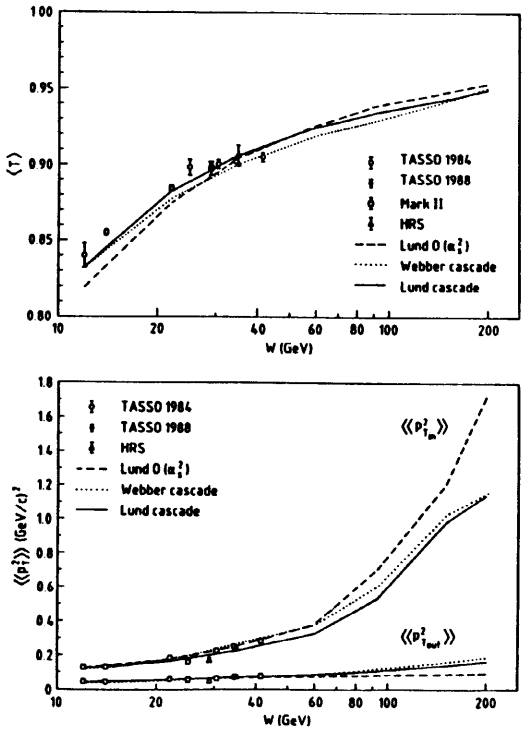

Figure 30: The evolution of average thrust and transverse momentum as a function of the hadronic CM energy W: TASSO data [3] compared with model results.

In [3] the W dependence of average event properties were analysed. They were compared to the models using the parameters optimized for $W = 35$ GeV. The general trend of the data is very well reproduced by all the models (see Fig. 30). Differences exist in detail. The measured average thrust between 20 and 45 GeV seems to change less than conceived in (all) the models. There is a slightly better agreement of the JETSET cascade option with the data. Discrepancies between model and data exist also for the average p_T in the event plane. Here the JETSET shower option predicts too small values. Webber and the second order JETSET option are in better agreement with the data.

No clear preference for either of these models can be inferred from the Q^2 dependence of the average values.

Some evidence for a stronger energy dependence of jet properties than anticipated in the second order formalism comes from the scaling violation of the fragmentation function [1]. The ratio R(x_p) with $x_p = 2p/W$ of the particle yield at $W = 14$ GeV to that at $W = 34$ or 44 GeV shows a significant drop of 25% for x_p >0.2 (Fig. 31). Whereas the second order matrix element calculation (using the traditional Q^2 scale, not the optimized scale discussed before) falls short of describing the size of the scaling violation, the shower approach provides a good description if the cut-off mass is set to the low value of 1 GeV.

Experiments at TRISTAN have reported on studies of jet properties on the basis

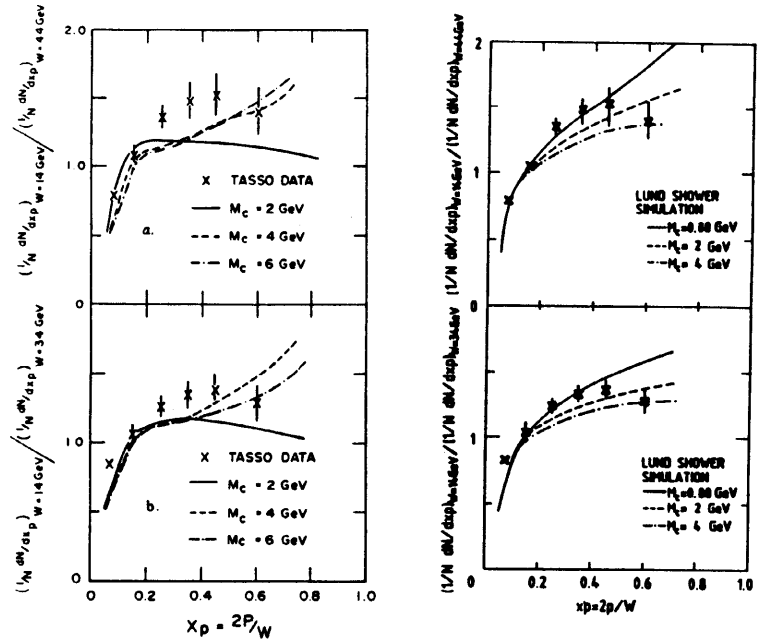

Figure 31: Scaling violation of the fragmentation function, for second order matrix element in JETSET (left), and for the shower algorithm in JETSET (right); results for different cut-off values compared with TASSO data [1].

of about a thousand hadronic events up to now (see e.g. [10]. Distributions like those of the cluster multiplicity, of the transverse momentum out of the event plane, or of the charged multiplicity were compared to the JETSET shower approach and the exact matrix element using their default parameters. Those are rather similar to the ones used by the Mark II Collaboration. In general the measurements can be well described. However, the data sample is too small for detailed conclusions.

A recent study on the energy dependence of n-jet rates may be found in [11]. While primarily directed towards perturbative QCD studies (like $\Lambda_{\overline{MS}}$ determinations), this kind of studies could be used to test existing models, not least parton shower based ones.

In summary: the global features of hadronic data can be described by fragmentation models based on QCD parton evolution and more or less ad hoc prescriptions for hadronisation. This is true at least at a fixed CM energy. Their predictive power for the evolution of the event properties with CM energy seems to be good, although there may be considerable deviations in detail. No systematic model comparison of the energy dependence of the complete particle and event distributions has been published up to now. As the lever arm in energy is increased by LEP, and later LEP 200, the energy dependence will provide increasingly stringent tests on models, see e.g. [12].

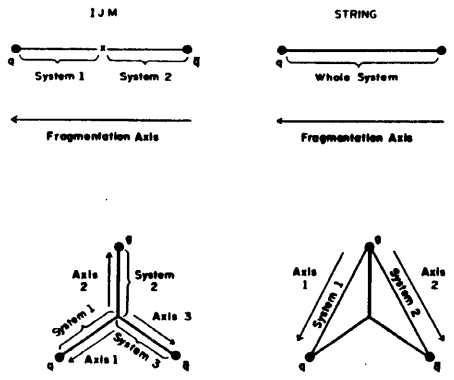

Figure 32: Schematic picture of differences between independent fragmentation (here denoted IJM for Independent Jet Model) and string fragmentation.

5.4 The String Effect

The particle and energy flow in the event plane has been shown to be a particularly good way for discriminating various approaches to hadronization. As was suggested in [13] it is a sensitive measure of the colour flow in the event.

It was applied to discriminate between the Lund (string) model and the independent fragmentation (a.k.a. independent jet) model. The latter was frequently used in the first years of data taking at PETRA and PEP. It views the hadronization process as developing independently for each parton along its direction. The prototype of these independent fragmentation models is that of Field and Feynman [14]. If supplemented by hard gluon bremsstrahlung [15] it can account for many features of the data. The basic concept of this model disagrees with the conjecture of colour confinement. Each quark or gluon fragments without any relation to a colour compensating partner. As a result the particle and energy flow in an event is distributed symmetrically around the individual partons.

In contrast to this, the hadronization occurs in the string model in colour neutral tubes formed between two partons.

It is easy to realize (see Fig. 32) that differences in the event topology predicted by these models should not show up if only two partons are produced (back to back). In both models the axis of hadronization coincides with the parton direction. If a gluon is emitted, this leads to an underlying three parton state. Now colour lines and parton directions do not fall on top of each other. In the string model they stretch between the quark and the gluon and the gluon and the antiquark (see Fig. 32). No direct colour connection exists between the quark and the antiquark. As a result the particle flow should fill the region between quark/antiquark and gluon leading to a depletion in the region between the quark and antiquark. This is different if all partons fragment independently from each other.

For the experimental tests [16] (see Fig. 33) three jet events were selected and the jets aligned in the event plane. The jet of highest energy ('jet 1') is directed along $\theta = 0$ degrees. The jet with the second largest energy ('jet 2'), in most cases a quark

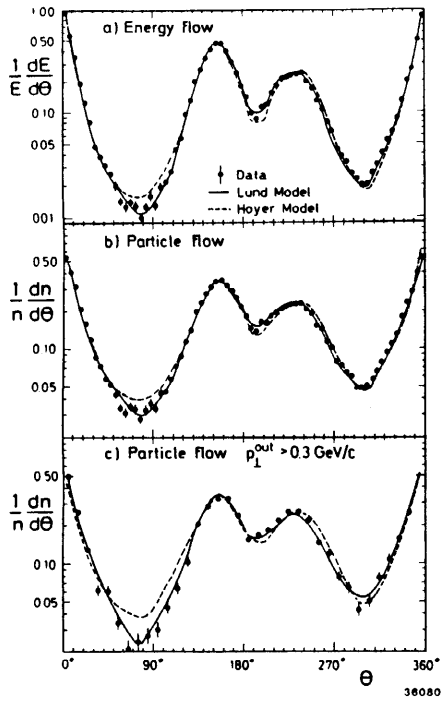

Figure 33: Energy and particle flow in event plane. JADE data [16] and model results.

or antiquark, is adjusted to have an angle smaller than π. The least energetic jet ('jet 3') which is supposed to be predominantly due to the gluon shows up at θ about 200 degrees. As can be seen from Figs. 33 and 34 the string model reproduces the measured particle and energy flow very well whereas the independent fragmentation model fails significantly. The failure is most prominent in the region between jet 1 and jet 2 reflecting the colour flow between quark and antiquark. The independent fragmentation model predicts too many particles in this region. The observed depletion is in agreement with the basic concept of the string model.

Additional and related measurements on the string effect have been presented. Notably the analysis of events where a photon instead of a gluon is emitted (suggested in [17] and experimentally tested in [18]) shows that the string effect is due to the different interplay between quark and gluons and quark/antiquark.

The energy and particle flow has also become a tool to study effects in the parton shower framework. As was calculated by [19] in the framework of the leading log expansion subsequent gluon emission is strongly correlated. The angle of emission of a gluon emitted at the nth branching has to be smaller than that originating from the $(n-1)^{th}$ branching. As a result the phase-space for low energy gluons becomes smaller leading to a depletion of the particle yield at low energies (rapidities). This was assumed to be the most prominent effect due to this coherence. Although the data agree with this conjecture [20] they are far from being conclusive: phase-space effects as well as

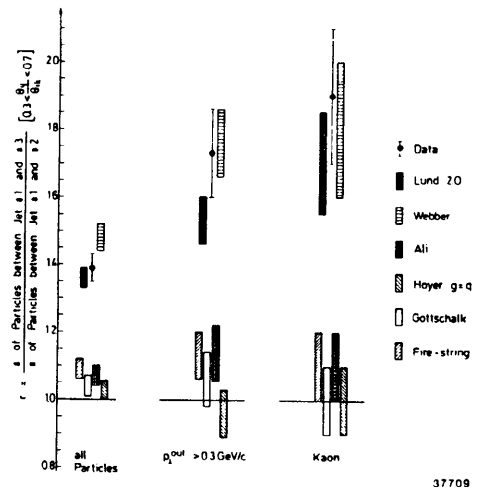

Figure 34: Ratio of particle flows jets (1-3)/(1-2). JADE data [16] and model results.

fragmentation effects tend to wash the dynamical effects out.

In a subsequent step this coherent gluon emission was incorporated into a Monte Carlo framework [6] and the energy and particle flow was tested. As a result the data could be very well described [21]. Shower models without coherence did not reproduce the string effect. In Fig. 34 the ratio of particle yields between jet 1 and jet 3 (mostly quark and gluon) to that between jet 1 and 2 (mostly quark and antiquark) is displayed as seen in the data and as predicted in various fragmentation models. Only the string model and the Webber model reproduce the data. The independent fragmentation models (Hoyer, Ali), a shower model without coherence (Gottschalk) and a model without hard gluon emission (Firestring [22]) predict too small a ratio. The theoretical evaluation of secondary gluon emission showed that the coherence leads to a colour flow which can be regarded as being due to two incoherent $q\bar{q}$ systems [17]. As can be expected from the discussion above, the string phenomenon does not necessarily require coherent gluon emission. In the shower formalism the observed energy and particle flow can be reproduced without coherent gluon branching if the hadronization is modelled along the string mechanism [23].

To summarize: the energy and particle flow is an effective tool to discriminate between various models of jet development. Both the string model and the shower model with coherent gluon emission can accomodate the observed features. The independent fragmentation model and shower models without coherence but using cluster decays fail. This is not the only distribution showing agreement between string models with fixed order QCD corrections and QCD shower models including coherence. The dip in the rapidity as well as the energy dependence of the maximum in the $\ln(x_p)$ distribution show very similar features. Again those cannot be reproduced by models based on independent fragmentation. In the technical terms of the simulation programs the string effect seems to be a result of a specific interplay between the generation of the underlying parton distribution and its subsequent transformation into hadrons.

5.5 Identified Particles

The analysis of the production of specific types of hadrons provides insight into the details of the local structure of the hadronization. Measurements have been performed on the total yield of various particles, on the dependence on kinematical variables, and on the correlations between particles of opposite flavour.

The experimental evaluation is complicated due to the additional requirements for particle identification leading to sizeable statistical and systematic errors. Still, a large amount of data has been collected at W's $\sim$ 30 GeV at PETRA and PEP and at W's $\sim$ 10 GeV at DORIS and CESR. Most pseudoscalar and vector mesons have been identified. In addition signals for 0^+ and 2^+ mesons ($f_0(975)$, and $f_2(1270), K_2^*(1430)$) have been seen. The considerable amount of baryons found within jets has provided special information about the fragmentation process. About 0.1-0.6 octet baryons (p, Λ, Ξ) have been found per event at CM energies around 30 GeV. Decuplet baryons are much less frequently produced. For many of those only upper limits could be given. It is common folklore that charmed and bottom hadrons can only be due to the primary interactions. They will be discussed separately.

Various ways of producing the variety of particle species have been implemented in the standard fragmentation models.

- In the traditional approach used in the independent fragmentation model and the string model $q\bar{q}$ pairs are successively created in the confinement colour field. They combine with the remnant quark antiquark from the primary interaction or the previous hadronization process to form hadrons. The probability of picking a certain quark flavour and of producing a hadron of a certain spin cannot be calculated but is given by some free parameters.
- The cluster decay is often applied as the hadronization step in shower models. Here the virtual gluons finally branch into $q\bar{q}$ pairs. Their flavours are given by some probability function. These quarks are then combined into colour neutral clusters which then decay according to phase-space.

As mentioned before, the number of free parameters differs considerably in the various models. For example the Lund model requires apart from the quark and diquark masses at least 5 parameters to steer meson and baryon production. Several other parameters are foreseen for more detailed considerations. The phase-space model conceptually requires none. However reality has turned out to be more complicated: to accomodate the measurements additional parameters had to be introduced. (Recently a string type model has been proposed [24] in which hadrons are produced according to their masses, i.e. with no explicit dependence on the masses of the constituent quarks is included, and therefore with few parameters. The model describes the particle yields well, but is not (yet) part of the standard models.)

In Fig. 35 the observed rates of the various particle species are displayed together with the predictions from various models (taken from [25]). The data are the averages from the PETRA and PEP experiments normalized to W = 29 GeV. In general the models describe the production yields quite well. The Lund model reproduces the measurements best. This is not surprising in view of the many parameters that can be adjusted. However, note that no prediction for tensor mesons existed in the version used for this comparison (provisions to generate them have only been included in the recent

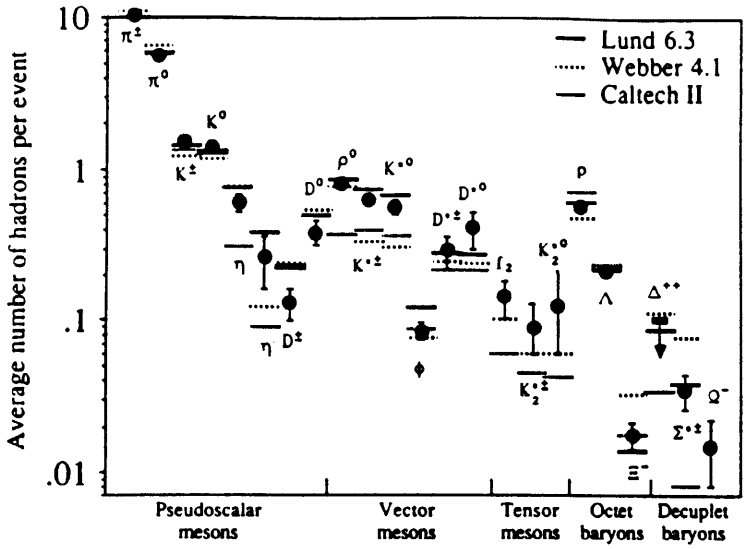

Figure 35: Average number of hadrons per event produced in e^+e^- annihilation at $\sqrt{s}$ = 29 GeV, averaged over PEP and PETRA experiments [25], and compared with models (default parameters).

version 7.1). The Webber simulation underestimates the production of strange mesons, most prominently for the K^*'s. In addition the baryon sector is badly described: the prediction exceeds the observed rate for heavier baryons (Ξ, Σ^*). The CALTECH-II model has a similar problem with the K^*'s but reproduces the baryon yield somewhat better. To parametrize the production of baryons most models adopt the notation of diquarks. Without them the relative yields would be determined by the folded probabilities of single $q\bar{q}$ pairs in the sea and by phase-space. Models like that of [26] based on this simple parametrization failed.

All models have difficulties in describing the measured rate of $0.014 \pm 0.006 \pm 0.004$ Ω's per event [27] . It is much larger than expected. The Lund model expects about two orders of magnitude less. The Ω yield is even larger than the 90% confidence upper limit of Ξ^{*-} production (0.008 per event). This is surprising since the Ω contains an additional s quark and is therefore expected to be subject to a further suppression. The assessment of this problem has to await more precise data. The one (published) measurement on Ω production requires confirmation, and observed rates instead of upper limits for the other decuplet baryons would help. (Ω production has also been observed at $W \sim 10$ GeV [28]. Better agreement with the expectation from Lund has been found at these low energies.)

Going one step further one can compare the fragmentation functions of the different particle species. Taking into account the feed-down from decays and the limited experimental accuracy of the data, all particles made up of light quarks have a similar fragmentation function and p_T dependence. A recent analysis [29] of data on the proton and Λ fragmentation functions suggests them being softer than those for mesons. This

is expected if diquarks are not pointlike.

Apart from studying the flavour dependence of particle production, the use of particles of known quantum numbers allows one to analyze how flavours are compensated within a jet. To do this one has to proceed from one particle distributions to two particle correlations.

In a first step experiments have studied the mechanism of charge compensation. Particularly two effects have been found.

- The charge of high momentum (rapidity) particles in one jet tends to be compensated by the charge of a high momentum (rapidity) particle in the opposite jet. This effect reflects the high Q^2 production of the primary quarks which become constituents of hadrons of high momentum, flying apart into opposite directions. This effect is reproduced by all models based on primary quark production and recursive emission of hadrons.
- At very low Q^2 Bose-Einstein symmetry leads to an excess of pairs of particles of the same charge over particle pairs of opposite charge [30]. This effect is usually neglected in fragmentation models, although models have been proposed [31]. It has only recently been provided for in the JETSET framework.

More information from two particle correlations can be extracted if the compensation of less frequently produced quantum numbers is studied. In particular the relation between a baryon and an antibaryon within a hadronic event shows the tendency of short range quantum number compensation. This contradicts models based on a stochastic hadron production as conceived in some models (for example [32]). It is reproduced by all of the standard fragmentation models. They are based on subsequent emissions of $q\bar{q}$ pairs.

The details of the short range behaviour provides a discriminative test for some hadronization mechanisms. In [33] the $p\bar{p}$ correlations were used in this respect. The angular distribution of a proton in the $B\bar{B}$ rest system was analysed and compared to the expectations from a cluster decay and from the string fragmentation. For the cluster decay the particles are expected to be isotropically distributed. In the case of the string fragmentation the particles are assumed to be produced with a limited p_T of 300 MeV but a considerably larger longitudinal momentum. As a result the protons/antiprotons tend to be aligned along the forward or backward direction in the $p\bar{p}$ rest system. Comparison with the data (see Fig. 36) shows a clear preference for the string model.

5.6 Heavy Flavour Jets

Information on the production and properties of heavy flavours in e^+e^- annihilation largely stems from two types of analysis: the study of inclusive lepton production [34], and the reconstruction of heavy hadrons through their hadronic decay [35]. Since prompt leptons arise from the semi-leptonic decays of both charmed and bottom flavoured hadrons, they provide insight, albeit indirectly, to the production mechanisms involved in both $c\bar{c}$ and $b\bar{b}$ events. Until recently, this had been the only way to study the hard fragmentation of bottom quarks. Of particular interest to LEP experiments, is, however, the progress made, with the aid of a vertex detection system, in tagging b events through the long decay length of the b-flavoured hadron [36]. The subsequent study of the rapidity and charged multiplicity distributions of the b-enriched sample

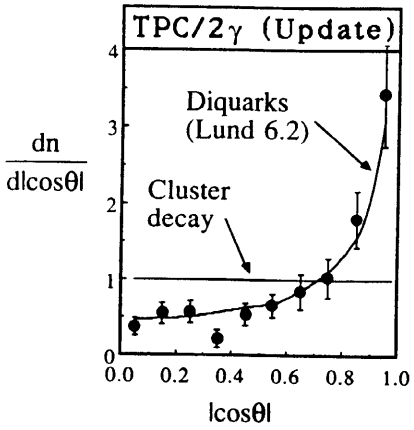

Figure 36: Distribution of the polar angle of the proton (antiproton) in the $p\bar{p}$ rest frame; TPC/2γ data [33] and model results.

unveils results that are in gratifying agreement with those from inclusive-lepton studies. They imply that $\langle x_E \rangle_b = E_{B\ hadron}/E_{beam}$ at 29 GeV is 0.76 ± 0.03, where the softening of the spectrum due to QED radiation has been corrected for. A comparison of the parton shower models shows that this is well represented by the Webber model ($\langle x_E \rangle_b \approx 0.73$), though the prediction of the Lund model ($\langle x_E \rangle_b \approx 0.88$) gives a spectrum that is evidently too hard. Replacing the default parameters of the Lund Monte Carlo with those suggested by TASSO [3] does not result in an improvement.

The method of reconstruction, on the other hand, allows a more direct analysis of the fragmentation of the heavy quark. It has been applied to a variety of charmed hadrons at PEP and PETRA energies, with notable success for the D^* meson. Since the majority of the reconstructed D^*s are the direct product of the transition of the primary charm quark to the primary hadron, the resulting energy spectrum closely corresponds to the charm fragmentation function. Other processes yielding D^*s, such as decays from higher charmed states and bottom hadrons, can largely be avoided at PEP and PETRA energies by studying only that part of the energy spectrum with $x_E = E_{D^*}/E_{Beam}$ greater than 0.5. At LEP energies, a harder cut may be necessary due to the relative increase (decrease) in the bottom (charm) cross-section. The production of charm within the shower itself is also realised at a level of $\approx 2\%$ of events. Again, it is pleasing to note that results from D^* analyses are in agreement with those obtained from prompt lepton studies. They give $\langle x_E \rangle_c = 0.57 \pm 0.03$ at 29 GeV, where the softening due to initial state radiation has been corrected for. A comparison of the parton shower models shows both Webber ($\langle x_E \rangle_c \approx 0.54$) and Lund ($\langle x_E \rangle_c \approx 0.62$) to be in agreement with data, though at opposite ends of the allowed band. Replacing the Lund parameters with those suggested by TASSO, results in a better agreement with data: $\langle x_E \rangle_c \approx 0.59$.

So what of the heavy hadron energy spectrum at the Z^0 pole? If we extrapolate the existing results on both charm and bottom data, to a centre of mass energy of 92 GeV, using the JETSET parton shower algorithmn (with parameter values given by TASSO) to describe the evolution of gluon radiation with increasing centre of mass, we would expect values of $\langle x_E \rangle_c = 0.48 \pm 0.04$ and $\langle x_E \rangle_b = 0.64 \pm 0.04$. These numbers can

be compared with the direct predictions of the Lund (Webber) parton shower Monte Carlo: $\langle x_E \rangle_c \approx 0.52$ (0.41) and $\langle x_E \rangle_b \approx 0.78$ (0.62). The discrepencies between the models at 29 GeV are likewise evident at 92 GeV. They should be clarified with the expected large yield of Z^0 events. More details together with the relative heavy particle production rates predicted by the shower models are to be found in an accompaning report [37].

5.7 Coherence Effects at LEP

Progress has been made in writing Monte Carlo simulation programs. Many important ingredients of perturbative QCD are nowadays taken into account. In particular, by the choice of an appropriate evolution parameter in parton showers (jet opening angle), it is possible to account for the angular dependence of soft emission, while maintaining a probabilistic interpretation. Reasonable fragmentation models are also available. The success of the resulting picture should not blind us to its shortcomings. Below, three examples are given, where naive notions do not agree with the expectations from perturbative QCD [38].

At PETRA/PEP energies, observed coherence phenomena could be explained entirely in terms of first order matrix elements (i.e. $q\bar{q}$ and $q\bar{q}g$ parton configurations) combined with the string fragmentation picture, without any reference to (perturbative) soft gluon emission. This will no more be true at LEP energies. Consider e.g. Figs. 33 and 34, where it was shown that string effects in the ratio of particle flows are enhanced by studying only particles with large momentum out of the event plane, p_T^{out}, or large hadron mass, m_h. If interpreted in terms of a Lorentz boosted string, this enhancement will be of the same strength at all CM energies, given fixed angles between jets. In the perturbative QCD scenario, on the other hand, the string effect is given by the number and energy of soft gluons emitted in the different interjet regions. No enhancement is then expected for the subsamples with large p_T^{out} or m_h [39]. While fragmentation effects may dominate at low energies, ultimately the perturbative QCD picture will win out. (This does not necessarily make the string concept wrong; it is only going to be irrelevant for these variables.) Fig. 37 shows the energy evolution of the ratios in the ARIADNE program, which contains both soft gluon emission and string fragmentation.

The canonical Lund scheme, in which three independent partons are replaced by two fragmenting string segments, in principle represented a step forward. However, even this scenario incorporates, in some sense, the traditional ideas of independent fragmentation, in that the two string segments are assumed to act as two uncorrelated sources of hadron production. The limitation of this approach to a basically quantum mechanical problem could be manifested, e.g., in double-inclusive correlations of the particle flow in three-jet events [41]. More specifically, one may compare the two ratios

$$\begin{aligned} r_1 &= \frac{dN}{d\Omega_{qg}} \Big/ \frac{dN}{d\Omega_{q\bar{q}}}, \\ r_2 &= \frac{dN}{d\Omega_{qg} d\Omega_{\bar{q}g}} \Big/ \frac{dN}{d\Omega_{q\bar{q}} d\Omega_{\bar{q}g}}, \end{aligned} \tag{157}$$

where $dN/d\Omega_{ab}$ represents the particle flow in the angular range between partons a and b. In the string scenario, the effect of the additional detection of particle flow in the $\bar{q}g$

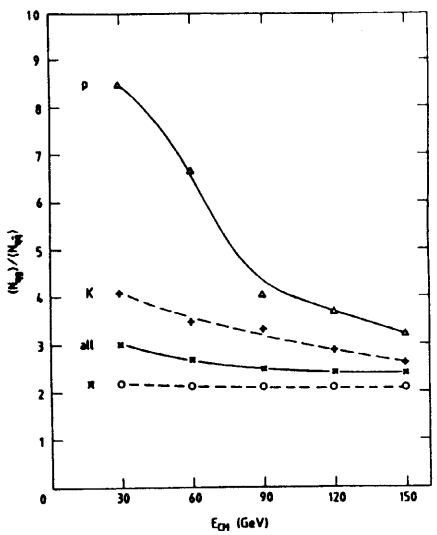

Figure 37: The ratio of the number of particles between the q and the g jet to that between the q and the $\overline{q}$ one, for symmetric three-jet configurations (using the central 40% of each angular range). Results for all particles, and separately for $\pi^{\pm}$, for $K^{\pm}$ and for p, $\overline{\text{p}}$, as obtained with the ARIADNE program at different CM energies [40].

sector cancels in the ratio of correlation functions (modulo fragmentation p_T effects). Therefore, in such a case, it is natural to expect $r_1 \simeq r_2$. Perturbative QCD, on the other hand, predicts $r_1 = 2.42$ and $r_2 = 2.06$ for symmetric three-jet events, i.e. an anticorrelation in the particle production between q and g and that between $\overline{q}$ and g. The relative smallness of this colour shielding effect does not diminish the fundamental importance of nonclassical effects.

Parton shower models, as well as string and cluster fragmentation models, are based on a neglect of colour suppressed terms, i.e. terms down in order by $1/N_C^2$ compared to the leading contribution. This is for a good reason: these terms often appear with a negative sign, or correspond to events with an undefined colour flow, and so we do not now know how to handle them in Monte Carlo simulation. Normally the neglect is not serious, but under special conditions $1/N_C^2$ terms may become sizeable or even dominant. Consider e.g. the emission of a soft gluon off a $q\overline{q}g$ configuration

$$\frac{dN}{d\Omega} \propto \frac{1-\overline{n}_q\overline{n}_g}{(1-\overline{n}_q\overline{n})(1-\overline{n}_g\overline{n})} + \frac{1-\overline{n}_{\overline{q}}\overline{n}_g}{(1-\overline{n}_{\overline{q}}\overline{n})(1-\overline{n}_g\overline{n})} - \frac{1}{N_C^2}\frac{1-\overline{n}_q\overline{n}_{\overline{q}}}{(1-\overline{n}_q\overline{n})(1-\overline{n}_{\overline{q}}\overline{n})} \tag{158}$$

with its well-know subdivision into a qg, a $\overline{q}g$ and a $q\overline{q}$ term, the last negative and neglected in shower programs. If the q and $\overline{q}$ are close to each other, this term is kinematically enhanced, and is no longer negligible. It can be studied experimentally by looking at the azimuthal distribution of particles around the quark direction, which is most conveniently done by tagging heavy flavour quark jets, see [42,38].

5.8 Intermittency

Unusually large density fluctuations in (pseudo)-rapidity have been observed in cosmic ray events [43] as well as in hadron-hadron [44,45], hadron-nucleus and nucleus-nucleus [46] collisions. These fluctuations have led to interpretations in terms of possible evi-

dence for a hadronic phase transition [47], hadronic Cerenkov radiation [48], hadronic hydrodynamics [49] or, simply, a cascading mechanism [50,51,52].

Before trying to distinguish between the different interpretations in terms of possibly new physics, the density fluctuations have to be shown statistically significant and not reproducable by present fragmentation and hadronization models.

To prove statistical significance, Białas and Peschanski [50,51] have suggested to study the dependence of the scaled factorial moments of order i

$$\langle F_i \rangle = \frac{1}{\langle \overline{n}_m \rangle^i} \langle \frac{1}{M} \sum_{m=1}^{M} n_m(n_m-1)\cdots(n_m-i+1) \rangle$$
$$\text{with} \quad \langle \overline{n}_m \rangle = \langle \frac{1}{M} \sum_{m=1}^{M} n_m \rangle \tag{159}$$

on the size of the rapidity resolution δy. In (159), the original rapidity window Δy (of say 4 units) is divided into M bins of size $\delta y = \Delta y/M$. The multiplicity is n_m in bin $m(m = 1...M)$. The average $\langle\rangle$ is over all events in the sample. The authors show the following:

1. Saturation of $\langle F_i \rangle$ with decreasing resolution δy is expected if the fluctuations are purely statistical, or the correlation is of range $d_0 > \delta y$.
2. If non-statistical self similar fluctuations of many different sizes in rapidity exist, the δy dependence follows the power law

$$\langle F_i \rangle \propto (\Delta y/\delta y)^{f_i} \quad \text{with } f_i > 0. \tag{160}$$

This latter effect is refered to as 'intermittency'.

Using these scaled factorial moments, the authors of [50,51] show that the JACEE density fluctuations [43] indeed provide indication for being intermittent.

Recent evidence for intermittent behaviour in nucleonic *accelerator* experiments comes from the Krakow-Minneapolis-Louisiana Emulsion group for pEm at 200 and 800 GeV as well as for $O^{16}Em$ at 60 and 200 A GeV [53].

Surprisingly, intermittency has been shown by NA22 [54] to be even stronger in hadron-hadron collisions at comparable energy. The slopes f_i for $0.1 < \delta y < 1.0$ are compared to those obtained by the Emulsion group [53] in Fig. 38a. There is a clear increase of the slope f_i with increasing order i, for all types of collision and, in particular for the higher moments, the f_i are larger for the NA22 data than in the $O^{16}Em$ data at similar energy.

Whatever the mechanism that causes intermittency in hh, hA and AA collisions, *it also seems to be present in e^+e^- collisions.* Buschbeck, Lipa and Peschansi [55] have been able to convert small δy negative binomial multiplicity fits of the HRS Collaboration [56] into scaled factorial moments via the relations [57]

$$\langle F_2 \rangle = 1 + 1/k \quad \text{and} \quad \langle F_3 \rangle = 1 + 3/k + 2/k^2, \tag{161}$$

thus connecting to the cascade picture of Van Hove and Giovannini [58].

Extracting the slopes f_2 and f_3 from the distributions for e^+e^- data (with a 2-jet cut) of Fig. 39 and comparing to the hadronic ones in Fig. 38a, we conclude that intermittency *is at least as strong in the former.* This finding has now been confirmed by

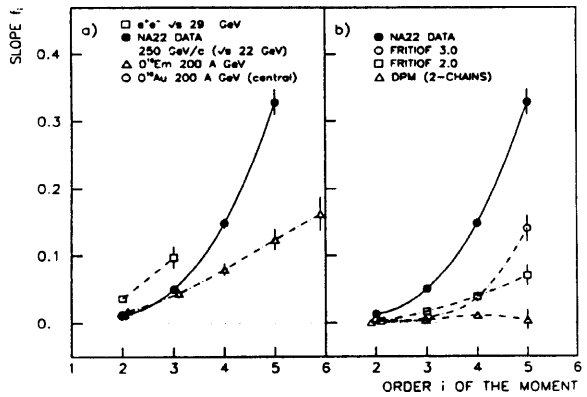

Figure 38: The slope f_i as a function of the order i for *(a)* the type of collision indicated, *(b)* NA22 events compared to DPM, FRITIOF2.0 and 3.0 [54].

a direct measurement in e^+e^- collisions by the TASSO Collaboration [59] and rules out the interpretation of intermittency as due to a hadronic reaction mechanism, including hadronic phase transition or hadronic Cerenkov radiation. It is, however, consistent with a cascade type fragmentation.

Can this intermittent behaviour be reproduced by currently used fragmentation and hadronization models? The fitted slopes of the two chain version of DPM [60] FRITIOF 2.0 and FRITIOF 3.0 [61] are compared to NA22 data in Fig. 38b. No intermittency can be expected from DPM, since this uses two (Lund) chains of independent particle and resonance production. Intermittency can, however, be expected in the FRITIOF model from the cascade-like treatment of multi-gluon radiation. It is better reproduced by version 3.0 than 2.0, but the effect is too low even there and something is missing. Furthermore, as shown in Fig. 39, the e^+e^- slopes at $\sqrt{s}$= 29 GeV/c are not reproduced by the cascade like JETSET shower model [62]. This conclusion has now been confirmed by the TASSO Collaboration [59].

The scaled factorial moment analysis according to eq. (159) is a standard service routine from JETSET 7.1 onwards [63]. In Fig. 40, $\langle F_2 \rangle$ is shown for the case of simultaneous subdivision in rapidity y with respect to the linear sphericity axis at the Z^0 peak ($\sqrt{s} = 92$ GeV). Four cases are considered as indicated in the figure. In the simple 2-jet case (not to be confused with the 2-jet *cut* of Fig. 39), the model even contains 'anti-intermittency', i.e. a suppression of nearby particle production. A separate slicing in y and ϕ (not shown) demonstrates that the effect is indeed in the latter, probably due to local p_T compensation. The use of second order matrix elements (mainly three-jet) gives larger factorial moments than the more realistic multijet configurations obtained with parton showers. The reason is the broadening of the jet profile in the showering. Bose-Einstein interference, as treated in this model by the routine LUBOEI, does not increase the effect considerably.

For a comparison of ARIADNE-3 and JETSET 6.3 to the HRS points, the normalisation in the analysis in JETSET was changed from that of [50] to that of [51] suited for all multiplicities. The conclusion of [55] is confirmed by a number of runs, both with and without Bose-Einstein interference included by means of the routine LUBOEI. In

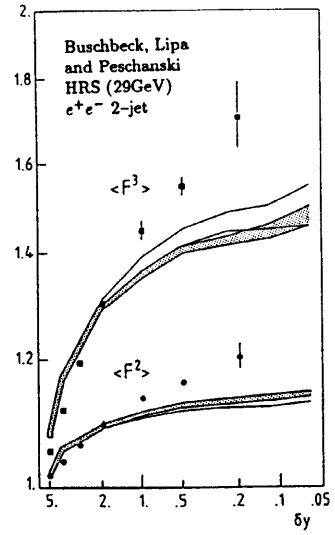

Figure 39: $\langle F_2 \rangle$ and $\langle F_3 \rangle$ deduced from the HRS 2-jet sample according to eq. (161), compared to calculations with the JETSET shower model. The shaded region corresponds to eq. (161), the non-shaded region to a direct calculation according to eq. (159) in the same δy region [55].

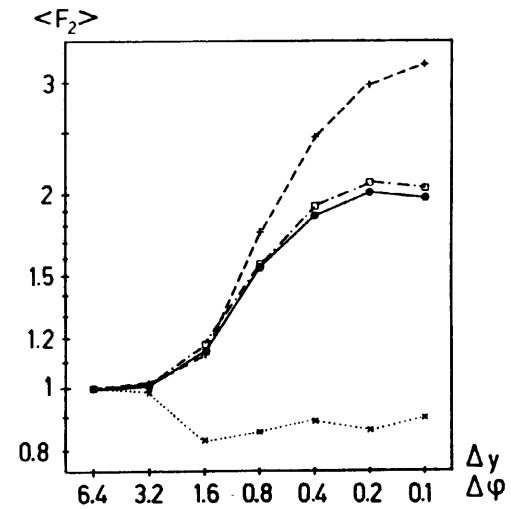

Figure 40: $\langle F_2 \rangle$ for e^+e^- events at the Z^0 peak in JETSET 6.3, as a function of the bin size ($\Delta y = \Delta\phi$). The dotted line is for 2-jet events, the dashed one with second order matrix elements. The full line corresponds to parton showers, the dot-dashed one with Bose-Einstein effects [63].

Fig. 41, the results of a standard parameter JETSET 6.3 run of 100 000 events, with Bose-Einstein effects, are compared to the $\langle F_i \rangle$ for the inclusive e^+e^- results from HRS [56]. Even though agreement looks better than for the 2-jet cut in Fig. 39, the model predictions are lower than the data, in particular for $\langle F_2 \rangle$ where the statistics is better, and become flat for δy smaller than the conventional short range order.

Also shown in Fig. 41 is a run of 100 000 events with ARIADNE-3, again including Bose-Einstein effects. Also this Monte Carlo comes close to the data, but does not show the increase for small δy. A likely explanation of the apparent failure of the FRITIOF, JETSET shower and ARIADNE Monte Carlos seems to be the following. Because of the self-similar character of these cascading models, intermittency is in principle present on the parton level (see also [64]). Due to necessary cut-offs and/or resonance decays, the effect is, however, smeared out over the conventional short range order during hadronization. Therefore, an improved hadronization picture seems to be needed and a generalized local parton-hadron duality [65,66] where $1/k$ (and via these $\langle F_i \rangle$) is preserved may be a good candidate.

We conclude that

1. *An increase of the normalized factorial moments with decreasing resolution down to $\delta y = 0.1$ (intermittent behaviour) is observed in all types of collision.*

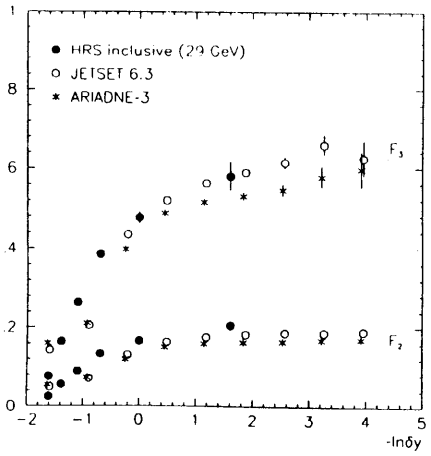

Figure 41: $\langle F_2 \rangle$ and $\langle F_3 \rangle$ deduced from the HRS inclusive sample according to eq. (161), compared to a direct calculation in JETSET 6.3 and ARIADNE 3.0 Monte Carlos including Bose-Einstein interference.

2. *The effect tends to be larger in e^+e^- and hadron-hadron than in nucleus-nucleus collisions.*
3. *This observation favours jet cascading over a hadronic reaction mechanism as possible interpretation, but presently used cascade models fail to reproduce the effect.*

5.9 Other Unresolved Issues

5.9.1 The e/π Ratio at Low Transverse Momentum

That electron production at low p_T is more abundant than can be expected from known sources (hadronic bremsstrahlung or decay of charmed particles) has been observed in several experiments [67] during the last decade. Furthermore, a dilepton continuum of mass $m < 0.6$ GeV has been observed both in the e^+e^- [68] and $\mu^+\mu^-$ [69] system, up to two orders of magnitude above that expected from the Drell-Yan process.

A possible explanation is 'soft annihilation' [70], where low mass lepton pairs are created through the annihilation of quarks and antiquarks produced during the hadronic collision. An alternative explanation has been suggested in terms of a thermodynamic model [71].

Of particular interest in this connection is the correlation between lepton and pion production. If the leptons are created *after* the final hadrons have been produced (i.e. by hadronic bremsstrahlung or hadron decays), the mean number of leptons per event will be proportional to the number of final hadrons, thus giving a constant e/π ratio. If, on the other hand, lepton pairs are produced over an extended volume at an early time *during* the collision, when new quarks and antiquarks have been created, lepton production is enhanced by interactions between the quarks and antiquarks produced in the hadron-hadron collision. The production rate then should be proportional to the density of quarks times that of antiquarks, i.e. to the square of the charged particle multiplicity [70,71]. In this case, the ratio e/π should increase linearly with the charged multiplicity n.

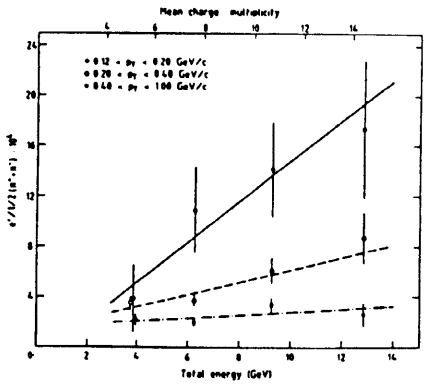

Figure 42: The residual e^+/π ratio after background subtraction as a function of total energy. The top scale indicates the charged multiplicity corresponding to the E_{tot} given on the bottom scale [72].

The production of prompt positrons has recently been studied [72] as a function of the associated charged particle multiplicity in the rapidity region $|y| < 1$. Fig. 42 shows the e^+/π ratio after background subtraction and efficiency corrections versus the central energy (depending linearly on the charge multiplicity, see upper scale), for three different p_T regions. While the e^+/π ratio is approximately constant for $p_T > 0.4$ GeV/c, it definitely increases with n at smaller p_T values.

This observation is important by itself, but even more when related to heavy ion collisions. It has been suggested that the expected phase transition between hadronic and quark matter could be detectable by just this quadratic multiplicity dependence of the dilepton production rate (last ref. [71])!

5.9.2 Soft Photons

Sea-quark annihilation [70] into virtual photons (observed as lepton pairs as discussed above) would imply also an excess of soft real photons. The study of soft photon production and its comparison to the electron-pair production would, therefore, be an ideal next step.

An excess of soft photons at $p_T < 60$ MeV/c of 30% above expectations from decay products of known particles and 5 times higher than expected from hadronic bremsstrahlung has indeed been found in K^+p collisions at $\sqrt{s} = 12$ GeV [73] and (earlier, but less significant) in π^+p collisions at 10.5 GeV [74].

More recently, the AFS Collaboration [75] has looked for a possible excess of photons, both in e^+e^- pairs from conversions and in electromagnetic showers. The results are not very conclusive, but in any case exclude a strong increase of the direct γ signal between $\sqrt{s} = 12$ and 63 GeV. However, the HELIOS Collaboration [76] finds a signal at $p_T < 30$ MeV/c in $p(Be, Al)$ collisions at 450 GeV/c, approximately 5 times higher than the level of hadronic bremsstrahlung, thus beautifully confirming the K^+p result. In the same analysis, the production of real photons is compared to that of virtual photons, i.e. lepton pairs, in terms of $M_T = \sqrt{M^2 + p_T^2}$. As shown in Fig. 43, the data seem to follow the same trend and thus suggest a common source.

Further evidence for the same underlying physics process is provided by the same

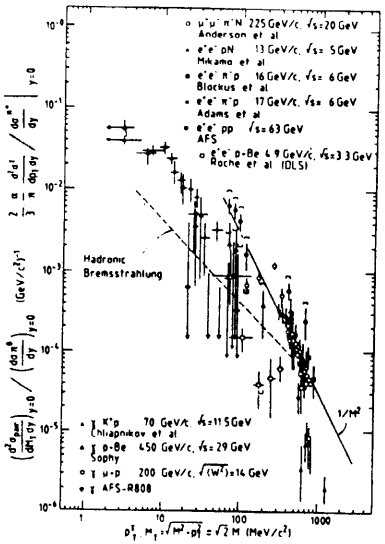

Figure 43: Compilation of low mass dilepton and real photon data plotted as a function of M_T and p_T, respectively. All data are normalized to the central pion multiplicity [76].

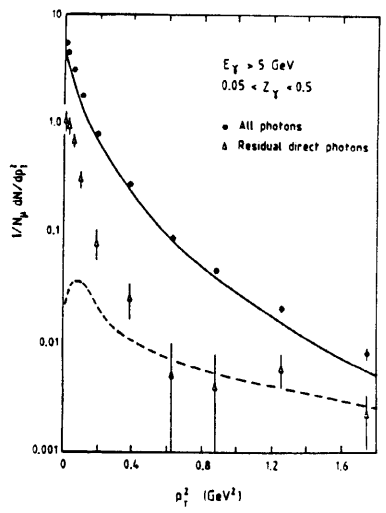

Figure 44: Normalized p_T distribution of all photons observed in the calorimeter. The continuous and dashed curves show the predicted contributions from hadron electromagnetic decays and muon bremsstrahlung, respectively, calculated in a complete simulation of the apparatus. The residual photons shown are determined from subtraction of the hadronic background only [77].

quadratic multiplicity dependence as that shown in Fig. 42, both in slope and absolute magnitude.

The effect is, however, not only present in hadronic collisions. As shown in Fig. 44 also the EMC Collaboration [77] has recently observed an excess of direct photons. The solid line corresponds to the Lund model prediction with parameters chosen to match the measured π^0 and ρ^0 inclusive yields in the same experiment. The Monte Carlo includes photons from all known hadronic electromagnetic decays. The excess photons are shown as triangles in Fig. 44 and, at $p_T^2 < 0.5(GeV/c)^2$, cannot be explained by an addition of muon bremsstrahlung (dashed) alone.

The experiment further shows that the prompt photons are emitted preferably along the struck quark direction. This suggests that their origin is at the quark level, but the multiplicity is an order of magnitude larger than expected from a radiating fractionally charged struck quark of conventional mass.

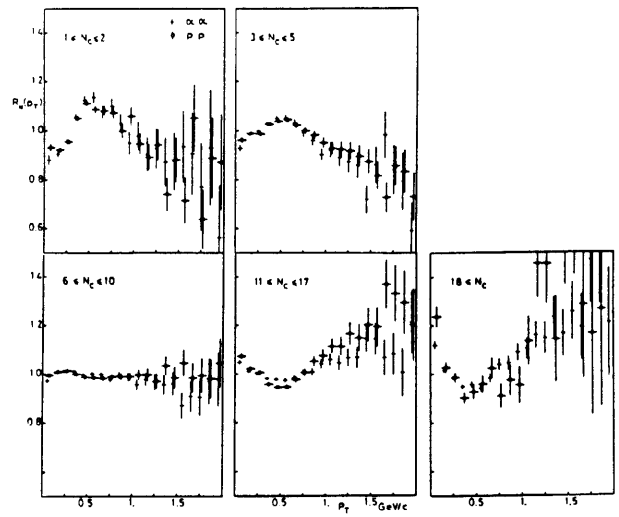

Figure 45: Normalized ratio of the p_T distributions for different multiplicity bins for $\alpha\alpha$ collisions at 31.2 GeV and pp collisions at 63 GeV [78].

5.9.3 Very Low p_T Hadrons

In Fig. 45, the normalized semi-inclusive p_T distributions

$$R_n(p_T) = (n^{-1}dn/dp_T)_n/((\langle n\rangle^{-1}\langle dn/dp_T\rangle) \tag{162}$$

are shown for charged particles produced in the central region ($|y| < 2$) of $\alpha\alpha$ collisions at $\sqrt{s_{NN}} = 31.2$ GeV and pp collisions at $\sqrt{s} = 63$ GeV [78]. The distributions are shown for different bins of the multiplicity. The shape of the function stongly varies with increasing multiplicity. For low multiplicity a dip is seen at small p_T and a bump at $p_T \approx 0.4$ GeV/c. On the contrary, at large multiplicity a steep maximum builds up at $p_T \approx 0$ and a minimum at $p_T \approx 0.4$ GeV/c.

Fig. 46 shows the p_T spectra for negative particles observed in pW and ^{16}OW collisions at 200 A GeV [79], respectively ($|y| < 0.9$). The dashed line gives the π^- spectrum from a MC generator representing pp collisions at FNAL and ISR after including nuclear effects. For both proton and ^{16}O data, a deviation from the generator is observed for $p_T \leq 0.025$ GeV/c.

Furthermore,in Fig. 47 is shown the normalized ratio of the p_T distribution of negative particles produced in ^{16}OAu and pAu collisions at 200 A GeV [80] to that for pp collisions. An effect strongly reminiscent of that for the high multiplicities in Fig. 45 is found and cannot be reproduced by the Lund-FRITIOF model [81].

5.9.4 Relevance for e^+e^- Collisions

Why are these observations important for e^+e^- collisions? Recently, Van Hove [82] has suggested a picture of cold quark-gluon plasma production, where the QCD cascade does not necessarily stop at a certain Q^2 cut-off, but can continue down to very low gluon virtualities. This super-soft process cannot be calculated in QCD and is not

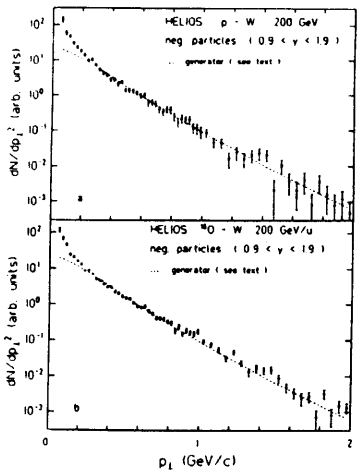

Figure 46: The p_T distribution of negative particles in pW and ^{16}OW collisions at 200 A GeV/c [79].

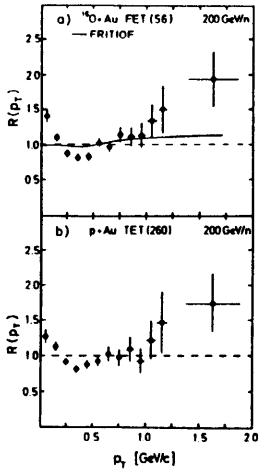

Figure 47: Ratio of the p_T distribution of negative particles in ^{16}OAu and pAu collisions at 200 A GeV/c to that of pp collisions [80].

included in present cascade Monte Carlo programs. Because it is the end of a gluon cascade process and not due to a (hot) quark-gluon plasma, it is indeed also expected in e^+e^- collisions. A possible evidence for that is the observation of soft photons in the EMC Experiment. The picture implies local parton-hadron duality [82,83,84,85] and may well alter our present ideas of parton hadronization.

References

[1] S.L. Wu, Phys. Rep. **107** (1984) 59;
B. Naroska Phys. Rep. **148** (1987) 67;
T. Sjöstrand, Int. J. Mod. Phys. **A3** (1988) 751;
P. Mättig, Phys. Rep. **177** (1989) 141

[2] Mark II Collaboration, A. Petersen *et. al.*, Phys. Rev. **D37** (1988) 1

[3] TASSO Collaboration, W. Braunschweig *et. al.*, Z. Physik **C41** (1988) 359;
P.N. Burrows, Ph.D. thesis, RAL-T-071 (1988)

[4] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rep. **97** (1983) 33

[5] T.D. Gottschalk, D. Morris CALT-68-1365 (1986)

[6] B.R. Webber, Nucl. Phys. **B238** (1984) 492

[7] JADE Collaboration, W. Bartel *et. al.*, Z. Physik **C33** (1986) 23;
S. Bethke, Habilitation thesis, LBL 50-208 (1987)

[8] TASSO Collaboration, W. Braunschweig *et. al.*, Phys. Lett. **B214** (1988) 286

[9] G. Kramer, B. Lampe, DESY 87-106;
N. Magnussen, Ph.D. thesis, DESY F22-89-01 (unpublished)

[10] A. Maki, Lecture given at the Banff Summer Institute 1988, KEK Preprint 88-52;
Y. Yamada, *in* Proceedings of the Third Lake Louise Winter Institute, Chateau Lake Louise, Canada, 6-12 March 1988, eds. B.A. Campbell, A.N. Kamal, F.C. Khanna, M.K. Sundaresan

[11] S. Bethke, LBL-26958 (1989)

[12] P.N. Burrows, Z. Phys. **C41** (1988) 375

[13] B. Andersson, G. Gustafson, T. Sjostrand, Phys. Lett. **B94**(1980) 211

[14] R.D. Field, R.P. Feynman, Nucl. Phys. **B161** (1978) 1

[15] P. Hoyer *et. al.*, Nucl. Phys. **B161** (1979) 349;
A. Ali, E. Pietarinen, G. Kramer, J. Willrodt, Phys. Lett. **B93** (1980) 155

[16] JADE Collaboration, W. Bartel *et. al.*, Phys. Lett. **B101** (1981) 129, Z. Physik **C21** (1983) 37

[17] Y. Azimov *et. al.*, Yad. Fiz. **43** (1986) 149, Phys. Lett. **B165** (1985) 147

[18] TPC/2γ Collaboration, H. Aihara *et. al.*, Phys. Rev. Lett. **57** (1986) 945;
Mark II Collaboration, P.D. Sheldon *et. al.*, Phys. Rev. Lett. **57** (1986) 1398

[19] A.H. Mueller, Phys. Lett. **B104** (1981) 161;
B.I. Ermolaev, V.S. Fadin, JETP Lett. **33** (1981) 269

[20] TASSO Collaboration, M. Althoff *et. al.*, Z. Physik **C22** (1984) 307

[21] JADE Collaboration, W. Bartel *et. al.*, Phys. Lett. **B134** (1985) 31;
TPC/2γ Collaboration, H. Aihara *et. al.*, Phys. Rev. Lett. **54** (1985) 270

[22] G. Preparata, G. Valenti, Nucl. Phys. **B183** (1981) 53

[23] M. Bengtsson, T. Sjöstrand, Nucl. Phys. **B289** (1987) 810

[24] C.D. Buchanan, S.B. Chun, Phys. Rev. Lett. **59** (1987) 1997, UCLA-HEP-89-003 (1989)

[25] W. Hofmann, Ann. Rev. Nucl. Part. Phys. **38** (1988) 279

[26] T. Meyer, Z. Physik **C12** (1982) 77

[27] Mark II Collaboration, S.R. Klein *et. al.*, Phys. Rev. Lett. **59** (1987) 2412

[28] ARGUS Collaboration, H. Albrecht *et. al.*, Phys. Lett. **B183** (1987) 419

[29] M.G. Bowler, P.N. Burrows, D.H. Saxon, Phys. Lett. **B221** (1989) 415

[30] TPC/2γ Collaboration, H. Aihara *et. al.*, Phys. Rev. **D31** (1985) 996;
CLEO Collaboration, P. Avery *et. al.*, Phys. Rev. **D32** (1985) 2294;
TASSO Collaboration, M. Althoff *et. al.*, Z. Physik **C30** (1986) 355;
for a review, see: B. Lörstad, Int. J. of Mod. Phys. **A4** (1989) 2861

[31] B. Andersson, W. Hofmann, Phys. Lett. **B169** (1986) 364;
M.G. Bowler, Z. Physik **C29** (1985) 617, Phys. Lett. **B180** (1986) 299, **B185** (1987) 205

[32] V. Cerny, P. Lichard, J. Pisut, Phys. Rev. **D16** (1977) 2822, Acta Phys. Pol. **B10** (1979) 629, Phys. Rev. **D18** (1978) 2409

[33] TPC/2γ Collaboration, H. Aihara *et. al.*, Phys. Rev. Lett. **57** (1986) 3140

[34] See for example: J. Chrin, Z. Physik **C36** (1987) 163

[35] See for example: S. Bethke, Z. Physik **C29** (1985) 175

[36] TASSO Collaboration, W. Braunschweig *et. al.*, Z. Physik **C42** (1989) 17

[37] See Report of the Heavy Flavour working group

[38] Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, DESY 88-093 (1988)

[39] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Phys. Lett. **B165** (1985) 147

[40] L. Lönnblad, private communication

[41] Yu.L. Dokshitzer, S.I. Troyan, V.A. Khoze, Yad. Fiz. **47** (1988) 1384 (translated Sov. J. Nucl. Phys. **47** (1988) 881)

[42] Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Leningrad preprint LNPI-1372 (1988)

[43] T.H. Burnett *et. al.* (JACEE), Phys. Rev. Lett. **50** (1983) 2062

[44] G.J. Alner *et. al.* (UA5), Phys. Rep. **154** (1987) 247

[45] M. Adamus *et. al.* (NA22), Phys. Lett. **B185** (1987) 200

[46] M.I. Adamovich *et. al.* (EMU-01), Phys. Lett. **B201** (1988) 397

[47] R.C. Hwa, Phys. Lett. **B201** (1988) 165, 'Multiparticle Fluctuation as a Possible Signature for Collective Effects', Oregon preprint OITS 385 (1988)

[48] I.M. Dremin, JETP Lett. **30** (1980) 152, Sov. J. Nucl. Phys. **33** (1981) 726

[49] J. Dias de Deus, Phys. Lett. **B194** (1987) 297

[50] A. Białas, R. Peschanski, Nucl. Phys. **B273** (1986) 703

[51] A. Białas, R. Peschanski, Nucl. Phys. **B308** (1988) 857

[52] W. Ochs, J. Wosiek, Phys. Lett. **B214** (1988) 617

[53] R. Hołyński *et. al.* (KML), Phys. Rev. Lett. **62** (1989) 733

[54] I.V. Ajinenko *et. al.* (NA22), Phys. Lett. **B222** (1989) 306

[55] B. Buschbeck, P. Lipa, R. Peschanski, Phys. Lett. **B215** (1988) 788

[56] M. Derrick *et. al.* (HRS), Phys. Lett. **B168** (1986) 299
and M. Derrick, ANL-HEP-LP-86-26

[57] A. Giovannini, Lett. Nuovo Cim. **6** (1973) 514

[58] A. Giovannini, L. Van Hove, Z. Physik **C30** (1986) 391, Acta Phys. Pol. **B19** (1988) 495

[59] A. Eskreys, Proc. Multiparticle Dynamics, La Thuile 1989, eds. A. Giovannini and W. Kittel (World Scientific), to be publ.

[60] A. Capella *et. al.*, Phys. Lett. **B81** (1979) 68, Z. Physik **C3** (1980) 329
M. Adamus *et. al.* (NA22), Z. Physik **C39** (1988) 311

[61] B. Andersson, G. Gustafson, B. Nilsson-Almqvist, Nucl. Phys. **B281** (1987) 289, 'A High Energy String Dynamics Model for Hadronic Interactions', Lund preprint LU TP 87-6

[62] T. Sjöstrand, M. Bengtsson, Comp. Phys. Com. **43** (1987) 367

[63] T. Sjöstrand, 'Multiple Interactions, Intermittency, and Other Studies', LU TP 88-20, Proc. Perugia Workshop on Multiparticle Dynamics 1988, to be published.

[64] B. Andersson, P. Dahlqvist, G. Gustafson, Phys. Lett. **B214** (1988) 604

[65] Ya.I. Azimov *et. al.*, Z. Physik **C27** (1985) 65, **C31** (1986) 213;
G. Cohen-Tannoudji, W. Ochs, Z. Physik **C39** (1988) 513

[66] L. Van Hove, A. Giovannini, Acta Phys. Pol. **B19** (1988) 917 and 931;
L. Van Hove, 'Cold Quark Gluon Plasma and Multiparticle Production', CERN TH5236/88, to be published in the Herman Feshbach Festschrift in Annals of Physics.

[67] L. Baum *et. al.*, Phys. Lett. **B60** (1976) 485;
E.W. Beier *et. al.*, Phys. Rev. Lett. **37** (1976) 1117;
M. Barone *et. al.*, Nucl. Phys. **B132** (1978) 29;
A. Maki *et. al.*, Phys. Lett. **B106** (1981) 423;
T. Åkesson *et. al.*, Phys. Lett. **B152** (1985) 411

[68] J. Ballam *et. al.*, Phys. Rev. Lett. **41** (1978) 1207;
S. Mikamo *et. al.*, Phys. Lett. **B106** (1981) 428;
D. Blockus *et. al.*, Nucl. Phys. **B201** (1982) 205;
M.R. Adams *et. al.*, Phys. Rev. **D27** (1983) 1977;
G. Roche *et. al.* (DLS), LBL preprint (1988)

[69] K.J. Anderson *et. al.*, Phys. Rev. Lett. **37** (1976) 799;
K. Bunnell *et. al.*, Phys. Rev. Lett. **40** (1978) 136;
M. Faessler *et. al.*, Phys. Rev. **D17** (1978) 689;
W.M. Morse *et. al.*, Phys. Rev. **D18** (1978) 3145;
J. Alspector *et. al.*, Phys. Lett. **B81** (1979) 397;
B. Haber *et. al.*, Phys. Rev. Lett. **D22** (1980) 2107;
R.I. Dzhelyadin *et. al.*, Nucl. Phys. **B179** (1981) 189

[70] V. Cerny, P. Lichard, J. Pisut, Acta Phys. Pol. **B9** (1978) 901, **B10** (1979) 537, Phys. Rev. **D24** (1981) 652, Z. Physik **C31** (1986) 163

[71] E.V. Shuryak, Phys. Rep. **61** (1980) 71;
T. Adachi, I. Yotsuyanagi, Phys. Rev. **D23** (1981) 1106;
R.C. Hwa, K. Kajantie, Phys. Rev. **D32** (1985) 1109

[72] T. Åkesson *et. al.* (AFS), Phys. Lett. **B192** (1987) 463

[73] P.V. Chliapnikov *et. al.*, Phys. Lett. **B141** (1984) 276

[74] A.T. Goshaw *et. al.*, Phys. Rev. Lett. **43** (1979) 1065

[75] Åkesson *et. al.* (AFS), 'Search for Direct γ Production at Low Transverse Momentum in 63 GeV pp Collisions', CERN-EP/87-85

[76] U. Goerlach (HELIOS), 'A New Measurement of Low p_T Direct Photons and a Comparison to the Low Mass Dilepton Continuum', Proc. XXIV Int. Conf. on High Energy Physics, Munich 1988, eds R. Kotthaus and J.H. Kühn (Springer 1989) p.1412

[77] J.J. Aubert *et. al.* (EMC), Phys. Lett. **B218** (1989) 248

[78] W. Bell *et. al.* (CHL), Z. Physik **C27** (1985) 191

[79] H.-W. Bartels (HELIOS), Z. Physik **C38** (1988) 85

[80] H. Stöbele *et. al.* (NA35), Z. Physik **C38** (1988) 89

[81] B. Andersson, G. Gustafson, B. Nilsson-Almqvist, Nucl. Phys. **B281** (1986) 289;
B. Nilsson-Almqvist, E. Stenlund, Lund preprint LU-TP 86-14 (1986)

[82] L. Van Hove, 'Cold Quark Gluon Plasma and Multiparticle Production', CERN TH5236/88, to be published in the Herman Feshbach Festschrift in Annals of Physics

[83] Ya.I. Azimov *et. al.*, Z. Physik **C27** (1985) 65, **C31** (1986) 213

[84] G. Cohen-Tannoudji, W. Ochs, Z. Physik **C39** (1988) 513

[85] B. Andersson, P. Dahlqvist, G. Gustafson, 'On Local Parton Hadron Duality, I. Multiplicity', LU TP 89-2

6 Standards

The writing of most event generators started in a University environment, as rather limited projects related to specific physics interests of the authors, and without too much thought of eventually making them available to experimentalists. Therefore, each author developed his own rules for things like particle code, event history and four-momentum representation. As the program grew, so did these conventions. For the outside user, getting acquainted with a new program therefore means learning a completely new set of rules. Furthermore, today we are in a situation where different programs are often good at different things, so that there is an ever-increasing need of interfacing generators with each other, and, of course, in the end, with detector simulation programs.

In recent years, the idea of setting common standards for program builders has therefore surfaced. In particular, the Particle Data Group spearheaded an effort to come up with a common standard for representing particle codes. This standard is now published [1], and has already appeared in event generators like EURODEC, HERWIG and JETSET. A brief summary is found in section 7.1. In the future, we expect this standard to be universally used.

Within our working group, we have tried to take the next step on this path, and devise a standard commonblock, in which events are to be stored. In the end, we came up with two commonblocks, where the second one is devoted to spin information, which is still not used by many programs, and where we also know of no solution that covers all eventualities.

It would be a major work to rewrite existing programs to use these commonblocks internally. More importantly, many programs need to use special kinds of information that could not easily be standardized (at least not before the day we have solved QCD, and *know* what the 'correct' approach to perturbative QCD and fragmentation is). We therefore propose the more modest solution, where each program builder keeps the commonblocks he already has, but adds an interface for translation to the standard commonblock. In its basic form, this would be a routine called after an event has been generated in full, and only requiring translation to the standard commonblock. To take full advantage of this approach, it would be necessary to translate also from the standard storing to the program-specific one. This is not always unique, but a rule of 'most reasonable interpretation' would have to be implied. In particular, with breakpoints suitably defined in the program, it should ultimately be possible to generate the hard process (including initial state radiation for e^+e^- events) in one program, the parton shower in a second, the fragmentation in a third, the decay in a fourth, perform detector simulation and reconstruction in a fifth, and analyze the event in a sixth.

The proposed standard was sent to most QCD event generator authors, and also to several experimentalists outside the LEP community. Responses were obtained from H.-U. Bengtsson, R. Brun, P. Burrows, F. Carminati, O. Di Rosa, T. Hansl-Kozanecka, G. Ingelman, S. Jadach, K. Kato, R. Odorico, F. Paige, J. Ranft, B. Ward, Z. Wąs and B.R. Webber. Suggestions for improvements and changes were discussed, and some modifications were introduced, while others were rejected. This led to a new draft, included in section 7.2. We have received enough encouragement, in particular from the program authors, that we hope this may now be considered a new *de facto* standard.

Standards are not only important for comparisons and interfacing of programs, but also for comparisons of data. Here we have not tried to give definitive answers, but

Table 15: Some of the most frequently used PDG standard particle codes [1]. Codes for antiparticles to the ones listed here, where existing, are obtained by adding a - sign. (Kindly note the misprint for the Ω^- code in the PDG listing; it is given correctly here.)

1	d	11	e^-	21	g
2	u	12	ν_e	22	γ
3	s	13	μ^-	23	Z^0
4	c	14	ν_μ	24	W^+
5	b	15	τ^-	25	H^0
6	t	16	ν_τ		
211	π^+	411	D^+	111	π^0
311	K^0	421	D^0	221	η
321	K^+	431	D_s^+	331	η'
130	K_L^0	310	K_S^0		
213	ρ^+	413	D^{*+}	113	ρ^0
313	K^{*0}	423	D^{*0}	223	ω
323	K^{*+}	433	D_s^{*+}	333	ϕ
2112	n	3112	Σ^-	3312	Ξ^-
2212	p	3212	Σ^0	3322	Ξ^0
3122	Λ^0	3222	Σ^+		
4122	Λ_c^+				
1114	Δ^-	3114	Σ^{*-}	3314	Ξ^{*-}
2114	Δ^0	3214	Σ^{*0}	3324	Ξ^{*0}
2214	Δ^+	3224	Σ^{*+}	3334	Ω^-
2224	Δ^{++}				
1103	dd_1	2101	ud_0	2103	ud_1
2203	uu_1	3101	sd_0	3103	sd_1
3303	ss_1	3201	su_0	3203	su_1

section 6.3 contains a few comments on current practices, which may be useful as a guideline. In particular, one should note that even as simple a property as the charged multiplicity may be (and has been) defined in slightly different ways.

Finally, note that there are also standards set e.g. by machine builders. One such is the coordinate frame of the incoming e^+ and e^- beams. Here the e^+ is taken to be moving in the $-z$ direction and the e^- in the $+z$ one. Of course, Monte Carlo authors should respect that.

6.1 The PDG Particle Codes

The standard particle codes are described in [1]. Some of the most frequently used codes are collected in Table 15.

A few general comments on the underlying principles:

- Where particle-antiparticle pairs exists, the antiparticle is distinguished by a - sign. The convention for which is particle and which antiparticle was settled by the PDG a few years ago; for the non-trivial case, see the note on mesons below.
- 'Fundamental' objects are given codes in the range below 100. In particular, quarks are in the range 1 – 10, leptons in the range 11 – 20, and standard model gauge bosons in the range 21 – 30. Quarks and leptons are arranged according to generation and, within each generation, according to weak isospin. Although not part of the official standard, it is generally foreseen to use 31 – 40 for non-standard model gauge bosons (H^+, Z'^0, etc.), 41 – 80 for supersymmetric partners, and 81 – 100 for internal program purposes. Actually, codes 91 – 94 are assigned specific such functions in section 6.2.
- Composite objects, made out of quarks, are also given composite codes, where the flavour and spin content is readily visible. In general, the last digit is $2 \cdot s + 1$, where s is the spin of the object, the previous three digits give the quark content in descending order (using the quark codes 1 – 6 of Table 15), and any digits before that are used to distinguish higher multiplets with the same flavour and spin content. A few exceptions exist. The rules are given in somewhat more detail in the following.
- A meson is normally given a code of the form $100 \cdot |q_1| + 10 \cdot |q_2| + 2 \cdot s + 1$, where the two relevant quark codes of the meson are assumed to be arranged so that $|q_1| \geq |q_2|$. If $|q_1| \neq |q_2|$, the meson is the one where q_1 is u, $\overline{s}$, c, $\overline{b}$, or t. Flavour mixed states are marked by having the last digit 0, and giving the flavour content in ascending order for the long-lived member. So far, only the K_S^0–K_L^0 system is used this way.
- A baryon is normally given a code of the form $1000 \cdot q_1 + 100 \cdot q_2 + 10 \cdot q_3 + 2 \cdot s + 1$, where the three quark codes of the baryon are arranged in descending order, $q_1 \geq q_2 \geq q_3$. An exception is Λ-like spin 1/2 baryons, where the last two quark flavours are given in reversed order, to distinguish them from the corresponding Σ-like baryons.
- A diquark is given a code of the form $1000 \cdot q_1 + 100 \cdot q_2 + 2 \cdot s + 1$, where the two quark codes are given in descending order, $q_1 \geq q_2$.

6.2 A Proposed Standard Event Record

The proposed standard contains two commonblocks, with contents as described below.

The main commonblock:
The main commonblock, with information on particle content, event history, momenta and production vertices, is

```
      PARAMETER (NMXHEP=2000)
      COMMON/HEPEVT/NEVHEP,NHEP,ISTHEP(NMXHEP),IDHEP(NMXHEP),
     &JMOHEP(2,NMXHEP),JDAHEP(2,NMXHEP),PHEP(5,NMXHEP),VHEP(4,NMXHEP)
```

Here the different parameters, variables and arrays have the following meaning.

NMXHEP: maximum numbers of entries (partons/particles) that can be stored in the commonblock.

NEVHEP: is normally the event number, but may have special meaning, according to the description below.

≥ 1: event number, sequentially increased by 1 for each call to the main event generation routine, starting with 1 for the first event generated.
= 0: for a program which does not keep track of event numbers.
= -1: special initialization record, see note 3 below.
= -2: special final record, see note 4 below.

NHEP: the actual number of entries stored in current event. These are found in the first NHEP positions of the respective arrays below. Index IHEP, $1 \leq$IHEP$\leq$NHEP, is below used to denote a given entry.

ISTHEP(IHEP): status code for entry IHEP, with following possible values.
= 0: null entry.
= 1: an existing entry, which has not decayed or fragmented. This is the main class of entries which represents the 'final state' given by the generator.
= 2: an entry which has decayed or fragmented and therefore is not appearing in the final state, but is retained for event history information.
= 3: a documentation line, defined separately from the event history. This could include the two incoming reacting particles, etc.
= 4 - 10: undefined, but reserved for future standards.
= 11 - 200: at the disposal of each model builder for constructs specific to his program, but equivalent to a null line in the context of any other program.
= 201 - : at the disposal of users, in particular for event tracking in the detector.

IDHEP(IHEP): particle identity, according to the Particle Data Group standard. Four additional codes have been defined for use in making event history more legible, see note 5.

JMOHEP(1,IHEP): pointer to the position where the mother is stored. The value is 0 for initial entries.

JMOHEP(2,IHEP): pointer to position of second mother. Normally only one mother exists, in which case the value 0 is to be used.

JDAHEP(1,IHEP): pointer to the position of the first daughter. If an entry has not decayed, this is 0.

JDAHEP(2,IHEP): pointer to the position of the last daughter. If an entry has not decayed, this is 0. See also note 6.

PHEP(1,IHEP): momentum in the x direction, in GeV/c.
PHEP(2,IHEP): momentum in the y direction, in GeV/c.
PHEP(3,IHEP): momentum in the z direction, in GeV/c.
PHEP(4,IHEP): energy, in GeV.
PHEP(5,IHEP): mass, in GeV/c^2. For spacelike partons, it is permissible to use a negative mass, according to PHEP(5,IHEP) $= -\sqrt{-m^2}$.
VHEP(1,IHEP): production vertex x position, in mm.
VHEP(2,IHEP): production vertex y position, in mm.
VHEP(3,IHEP): production vertex z position, in mm.
VHEP(4,IHEP): production time, in mm/c ($= 3.33 \cdot 10^{-12}$ s).

The following notes have been added, with some clarifications and further definitions.

1. The three characters HEP (for High Energy Physics) have been chosen as a standard part of all names, in order to make them easily distinguishable and to avoid overlaps with other names.
2. The default size has been chosen to be 2000 entries, with the option to change that via the parameter construction. It is assumed that active use will be made of the NMXHEP value in the translation routine, to ensure that this limit is never surpassed.
3. A record with NEVHEP = -1 may optionally be written by an initialization routine. This record may be arbitrarily large, and contain program-specific information, but the first entry (IHEP = 1) is reserved for standard quantities:
   ISTHEP(1): identifier for run, defined by user.
   IDHEP(1): identifier for generator used, defined by user.
   JMOHEP(1,1): date of run, in form yymmdd.
   JMOHEP(2,1): time of run, in form hhmmss.
   JDAHEP(1,1): version and subversion of generator, in form vvss.
   JDAHEP(2,1): last date of change of generator, in form yymmdd.
   Additional lines may e.g. contain information on subprocesses switched on.
4. A record with NEVHEP = -2 may optionally be written by a routine called after events have been generated. This record may be arbitrarily large, and contain program-specific information, but the first entry (IHEP = 1) is reserved for standard quantities:
   ISTHEP(1): number of events generated in current run.
   PHEP(1,1): total cross-section for the simulated process(es), as obtained by Monte Carlo integration during the course of the event generation, in mb.
5. The following four codes have been defined, in addition to the standard ones given by the PDG:
   91: a cluster;
   92: a string;
   93: an independent fragmentation parton system; and
   94: a showering parton system.
   Thus a particle produced in fragmentation should not point back directly to the parton it comes from, which anyway (usually) is not well-defined, but rather point to one of the objects 91 – 93 above. Similarly 94 may be used to define the effective CM frame of a showering parton system.
   Note that the flavour content of this object, as well as details of momentum sharing, has to be found by looking at the mother partons, i.e. the two partons in positions JMOHEP(1,IHEP) and JMOHEP(2,IHEP) for a cluster or a shower system, and the range JMOHEP(1,IHEP) – JMOHEP(2,IHEP) for a string or an independent fragmentation parton system.
6. It is assumed that the daughters of a particle (or another object like a cluster or a string) are stored sequentially, so that the whole range JDAHEP(1,IHEP) – JDAHEP(2,IHEP) contains daughters (with one exception, see note 7). Even in decays with only one daughter (e.g. $K^0 \to K^0_S$) both values should be defined, to make for a uniform approach in terms of loop constructions.
7. Normally daughters are stored after mothers, but in backwards evolution of initial state radiation the opposite may appear, i.e. that mothers are found below the daughters they branch into. Also, the two daughters need then not appear one after the other, but may be separated in the event record.
8. Only production vertices are given. For a particle which decays, the decay vertex can be found by looking at the production vertex of its daughters.

9. Some programs may not provide all output or input, and should then leave the corresponding positions blank. Example: a program without vertex information should just put VHEP = 0 on output; it is then up to the user to read the program manual and check what information a given program does indeed provide.
10. When reading in from /HEPEVT/ to the program-specific commonblocks, minimal assumptions should always be made. This means that lines which have ISTHEP(IHEP) ≥ 11 should be considered empty lines in the same sense as those with ISTHEP(IHEP) = 0, that particle codes unknown to the program should just be skipped and the corresponding entries left inert by future physics calls, that a jet system that is to undergo string fragmentation is already ordered with partons sequentially along the string, etc. Each Monte Carlo programmer may here want to specify exactly what is achievable by a simple translation call, and what needs special editing in the internal commonblocks afterwards.

The polarization commonblock:
Polarization information has not been included in the commonblock above, since it is not used by many programs (yet), and since no simple approach can give the full answer for all possible spins.

For a spin half unstable particle, it is convenient to keep record of its spin polarization density matrix in the form of a vector.

The following commonblock is therefore supplied, which is supposed to parallel /HEPEVT/, entry by entry:

```
COMMON/HEPSPN/SHEP(4,NMXHEP)
```

For a given entry, SHEP gives the four-vector s, according to the following well-defined convention. In the rest frame of the fermion, $p = (m; 0, 0, 0)$, the three-vector part $\vec{s}$ is related to the spin density matrix ρ in the conventional way, i.e.

$$\rho = \frac{1}{2}\vec{\sigma} \cdot \vec{s} = \frac{1}{2}(\sigma_x s_x + \sigma_y s_y + \sigma_z s_z), \tag{163}$$

where the σ_i are the ordinary Pauli matrices. In this frame, we define the contravariant four-vector to be $s = (1; \vec{s})$. The 1 for the zeroth component is a matter of convention, but it makes a number of expressions simpler.

The important second part of the prescription is the following. The transformation properties of the spin four-vector s is taken to be the same as for the four-momentum p, i.e. the same boosts and rotations are to be applied to both. In particular, if we later on come back to the fermion rest frame (e.g. to simulate the decay), we automatically keep track of a possible change of the quantization axis (Wigner rotation).

It is neither necessary nor profitable to put any restriction on the modulus of s (in the rest frame), apart from the trivial one $|\vec{s}| \leq 1$. A pure spin state corresponds to $|\vec{s}| = 1$. Correlations between particles (say in a $\tau^+\tau^-$ pair) may be simulated by an appropriate choice of the spin vectors at the production vertex; different strategies are here possible, but need not be specified as part of the standard.

6.3 Conventions for Inclusive Distributions

To obtain the 'real' inclusive distributions of hadron production in e^+e^- collisions, most experiments at PETRA and PEP corrected their distributions according to

$$\frac{(dn/dx)_{true}}{N_{true}} = \frac{\frac{(dn/dx)_{generator}}{N_{gen}}}{\frac{(dn/dx)_{MCdetected}}{N_{MC}}} \frac{(dn/dx)_{observed}}{N_{obs}}, \tag{164}$$

where x is the quantity studied, the capital N stands for the number of events, and the small n denotes the number of particles or tracks. In addition,

- $(dn/dx)_{true}$ is the true distribution;,
- $(dn/dx)_{generator}$ is the original distribution as generated in the Monte Carlo;
- $(dN/dx)_{MCdetected}$ is the distribution generated in the Monte Carlo and subjected to all experimental effects; and
- $(dN/dx)_{observed}$ is the directly measured distribution.

Obviously the Monte Carlo distributions are to approximate the real world as close as possible, both for the generator and the simulation of the detector.

In order to understand the systematic errors of the correction procedure, it is recommended to compare at least two event generators for the correction factor calculation; the generators should be based on as different physics assumptions as possible while still in reasonable agreement with data.

There is some ambiguity which particles and events to choose as those contributing to the generated distribution. Experiments at previous e^+e^- colliders adapted some conventions which in part have historical reasons and which are by no means straightforward. Still, taking into account these prescriptions simplifies the comparisons of low and medium energy e^+e^- results (e.g. [2,3,4,5]) with the measurements around and above the Z^0.

Ambiguous areas are particularly what kind of particles to consider as 'generated', which particles (or partons) to consider for the determination of the true axis, and what to do with QED bremsstrahlung. We will outline the most common approaches to these points.

In principle, one would like to include only those particles that emanate directly from the interaction point, i.e. before decays. However, since it is by no means clear how many of the different types of resonances are produced in a jet event this would lead to a large model dependence. On the other hand, particles like the K_L^0 decay far outside the detector and therefore their decay products do not contribute to the measured particle yields and should not be included in the generated sample. The borderline between these extremes is defined to include decay products from particles with a lifetime of $\tau \leq 3 \cdot 10^{-10}$s in the sample of generated events. Those of a longer lifetime are regarded as stable. Particularly this implies that all K_S^0's and Λ's will be treated as decayed, irrespective of how far they travelled. On the other hand K_L^0's or π's etc. are treated as stable, even if they decay in the detector. One should be aware that the omission of the pions from K_S^0 decays reduces the average charges multiplicity by about 10%.

In many e^+e^- experiments at lower energies only the charged tracks were measured reliably. This led to additional choices for defining the 'true' event axis and related quantities like thrust or aplanarity. In general the 'generated axis' was calculated from

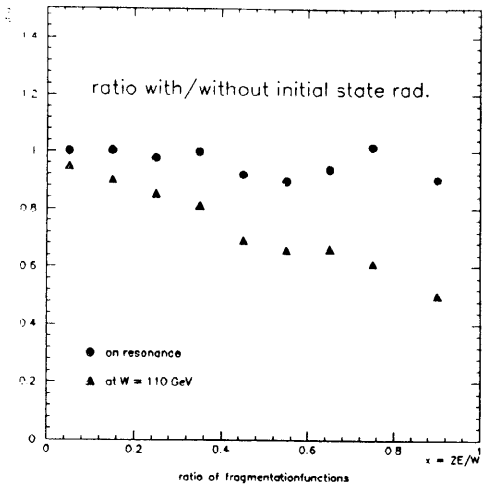

Figure 48: Ratios of momentum spectra, with and without initial state radiation, at 92 and 110 GeV CM energy.

the charged and neutral particles. These particles were put into the routines for calculating thrust and sphericity. On the other hand, the measured distributions, both on the Monte Carlo level and for the real data, were based on the charged tracks only. The corrected distribution still reflects the charged and neutral particles.

Of special importance for e^+e^- interactions is the QED bremsstrahlung. Photon radiation reduces the effective CM energy of the hadronic system leading to a range of effective CM energies. As a result the particle spectrum and topological event properties are changed. To correct for this effect, only events at the nominal CM energy are taken for the generated spectrum. On the Z^0 peak initial state radiation affects the particle spectrum only little. This is drastically different above the peak: the relevant distributions are quite strongly distorted. In Figure 48 the ratios of the momentum spectrum in terms of $x = 2 \cdot p/W$, with and without initial state radiation, are displayed for CM energies on resonance and at $W = 110$ GeV.

A very important step in understanding jet development is the analysis of the dependence of their properties on the CM energy. The measurements at the Z^0 will largely extend the lever arm for such an analysis. For a comparison to low energy data it is warmly recommended to apply the above prescriptions to the LEP data.

When comparing to multiplicity distributions in hadron-hadron collisions (e.g. [6,7,8]), the following differences have to be taken into account:

1. A distinction has to be made between diffraction-dissociation and non-(single-) diffractive mechanisms. Diffraction-dissociation in general leads to low multiplicities, but often is not excluded from the multiplicity distribution.

2. V^0 decays, Dalitz decays and γ conversions are in general excluded in hadron-hadron experiments, either by an extrapolation from decays observed at larger distances down to the vertex region [6], by Monte Carlo simulation [7] or by a proper renormalization [8].

3. Similarly, corrections are applied for secondary interactions of charged particles.

In recent studies [6,7,8], efficiencies and corrections are fed into a system of equations in the true and the observed multiplicities, and the system is solved for the true ones.

References

[1] Particle Data Group, G.P. Yost *et. al.*, Phys. Lett. **B204** (1988) 1 (codes are given on pp. 113-114)

[2] J. Siegrist *et. al.*, Phys. Rev. **D26** (1982) 969

[3] Mark II Collaboration, A. Petersen *et. al.*, Phys. Rev. **D37** (1988) 1

[4] TASSO Collaboration, M. Althoff *et. al.*, Z. Physik **C22** (1984) 307; TASSO Collaboration, W. Braunschweig *et. al.*, Z. Physik **C41** (1988) 359

[5] HRS Collaboration, D.Bender *et. al.*, Phys. Rev. **D31** (1985) 1

[6] M. Adamus *et. al.* (NA22), Z. Physik **C32** (1986) 475, Phys. Lett. **B177** (1986) 239, Z. Physik **C37** (1988) 215, Phys. Lett. **B205** (1988) 401

[7] K. Alpgård *et. al.* (UA5), Phys. Lett. **B121** (1983) 209; G.J. Alner *et. al.* (UA5), Phys. Lett. **B138** (1984) 304, **B160** (1985) 193, *ibid.* 199, **B167** (1986) 476; R.E. Ansorge *et. al.* (UA5), 'Charged Particle Multiplicity Distributions at 200 and 900 GeV c.m. Energy', CERN-EP/88-172; Ch. Fuglesang, 'A Method for Correcting Observed Distributions of Multiplicities Using the Maximum Entropy Principle', CERN-EP/88-174

[8] A. Breakstone *et. al.* (ABCDHW), Phys. Rev. **D30** (1984) 528 and 'Charged Multiplicity Distributions in Rapidity Bins for pp Collisions at $\sqrt{s}$ = 31, 44 and 62 GeV', CERN/EP 88-133

7 Summary and Recommendations

The preceding sections have concentrated on physics present (or, sometimes, not present) in programs. This may seem factual enough, but obviously some value judgment is always implicit in anything written (starting with the space devoted to the various programs, which was determined by a mixture of deliberate decisions and random factors). The summary and recommendations in this section are no exceptions to the rule. So, if you disagree with our conclusions, please feel free to do so. Remember, QCD is complicated enough that we do not know the true answer. It would therefore be dangerous to limit our frame of reference at a time like this, when LEP offers the opportunity to test, and maybe falsify, many of our preconceived notions.

There is a flip side to this coin. From most experimentalists point of view, the rôle of a QCD generator is to help them in understanding event structures and interpreting data. Unfortunately, for them, the situation is all too often the other way around, with data being used to test generators.

Having said that, let us now proceed with a set of observations, conclusions, and recommendations. The first conclusion is that no one single QCD program is best for all the subtasks of e^+e^- annihilation event generation. It may therefore be necessary to combine several programs for the best overall description. The subsequent discussion will be structured according to the main issues involved.

Hard cross-section and initial state radiation:
Most programs have the full Born level γ/Z^0 cross-section included. Modifications to this cross-section come mainly from three sources. In order of ascending importance, they are:

- Final state QCD corrections. These give an extra factor R_{QCD}, eq. (3). While the size of the third order term has caused some worry, see the QCD theory report, the net contribution is known to be small, and is easily included. In fact, many programs include R_{QCD}, at least to second order.
- Electroweak loop corrections. These are covered in detail in the electroweak theory reports. The full expressions, as known today, are very lengthy, and beyond the scope of any QCD generator. However, a very good approximation can be obtained by using an effective description, in terms of a modified Z^0 mass, width and coupling strength, and a modified $\sin^2\theta_W$ in the calculation of axial vector couplings of quarks to the Z^0. This machinery is not included in any of the QCD generators, but it would be no problem to do so. In some programs this only requires a judicious reinterpretation of already existing program parameters, i.e. no change at all in the actual code.
- Initial state QED radiation. Since the effect of initial state photon emission is to reduce the mass of the γ/Z^0 propagator, large changes of cross-sections are induced by this mechanism, in particular in the region around the Z^0 peak. Many QCD programs have nothing to offer here. JETSET, NLLJET and TIPTOP include emission of (at most) one photon, in lowest unexponentiated order. This is known to be insufficient in the Z^0 region. COJETS has been announced to include multiple photon emission, but this version is not yet ready for distribution, so we can say nothing about its quality.

In view of the last point above, in particular, we recommend that people do *not* use the existing QCD generators to describe initial state radiation and the subsequent interaction, at least not for any precision tests. Instead, dedicated electroweak generators should be used for this task. The field of possibilities is described in the electroweak Monte Carlo report. The task of generating a $q\bar{q}$ pair is rather similar to that of generating a $\mu^+\mu^-$ pair, but note that no muon pair generator can be used as it is. Not only have couplings and masses to be changed, and a $N_C R_{QCD}$ factor to be included, but furthermore final state QED radiation and initial–final state interference has to be switched off (in a consistent manner), for reasons to be discussed in the next section.

Final state QED radiation and interference:
In the final state, the original $q\bar{q}$ pair successively loses its initial energy, by gluon radiation (whether it be hard or part of the soft confinement potential). At some stage, the fragmentation process proper starts, leading to the formation of the final state hadrons. Thus, while hard, non-collinear final state photon emission should be describable in terms of the perturbative quark picture, and soft photon emission by the current of final state hadrons, there exists an uncomfortably large region in between where we do not know the correct procedure. This obviously not only affects the description of final state QED radiation, but also that of initial–final state interference.

Given the uncertain state of affairs, it is maybe not surprising that no existing program, neither QCD nor electroweak, addresses these issues. A first step would be to include hard, non-collinear photon emission as part of the shower evolution algorithm, with the cut-off of the photon emission as an additional parameter, independent of the standard shower cut-off for gluon emission. We recommend Monte Carlo authors to consider this seriously. The actual extra work for a crude implementation into an existing shower algorithm is modest, and yet it would provide some of the tools needed for an experimental study, which could pave the way for an eventual better understanding.

From a practical point of view, neither final state radiation nor initial–final interference are expected to have major implications for any of the planned LEP studies, so the issue is somewhat decoupled from other aspects.

Perturbative QCD:
Today, two approaches to perturbative QCD are available, the matrix element one and the parton shower one. Both have been extensively described. In weighting one against the other, we recommend parton showers as the standard option, to be used when there are no obvious counterarguments. The reason is that showers are expected to give the most realistic overall picture of the complex multijet structure of e^+e^- hadronic events at high energies. Applications for which matrix elements should nonetheless be used include α_S determinations, studies of the three-gluon vertex, and other similar topics.

Matrix elements:
Some QCD studies, like α_S determinations, require all matrix elements to be given to the same order, i.e. in practice second order. The main comments on second order matrix elements are as follows.

- The y cut-off approach is to be preferred over the (ϵ, δ) one, so as to keep the number of parameters and the amount of possible confusion to a minimum.
- Based on experience from lower energies, a low y cut-off value should be used;

if possible $y = 0.01$ is a good value. Even with this value, the implied invariant mass cut-off of $\simeq 9$ GeV is uncomfortably large, but small values are excluded, since then the two-jet rate would be negative.

- Several implementations exist which give rather different results. Unless one can explicitly show that some implementations are wrong, the disagreement should be taken as an indication that systematic method errors are not negligible, and properly include this in any quoted α_S values.
- It should be clearly stated which implementation is used in a study. For future reference, it is helpful if experimental plots for a given quantity not only include the data and the full Monte Carlo result, but also the parton level result from the generator.
- Among the implementations reviewed, we see no reason to favour one above the other on theoretical grounds, except that the GKS option of JETSET, where the matrix elements are acknowledged to be incomplete, is disfavoured.
- From the point of view of compactness, and hence transportability, the Zhu parametrization stands out. It could be a suitable standard to refer to. The way the Zhu algorithm has been used in the past, with weighted events, is not to be recommended, however. The JETSET approach here seems to be preferred, for its greater flexibility and ease of use.
- The concept of optimized perturbation theory, applied to second order matrix elements, is growing in general acceptance, and will certainly play a prominent rôle in perturbative QCD studies at LEP. It is only this way the matrix element approach yields a reasonable description of two-, three- and four-jet rates simultaneously. After the recent successes, optimized scale matrix elements have become more appealing also for other studies, although we still expect the lack of multiple soft gluon emission to be visible at LEP.
  As to the numerology of this approach, values of $y' = Q^2/s \simeq 0.002 - 0.005$ seem preferred, i.e. $Q \approx 5$ GeV at LEP energies. With this choice, one obtains a small Λ, ≈ 100 MeV, but a large α_S, ≈ 0.18. The interpretation of these values is rather less clear than with the standard ansatz $Q^2 = s$.

Going beyond second order, we are pleased to note that now three independent calculations exist of third order five-jet matrix elements, and that they have been checked against each other to agree. Furthermore, these calculations have been turned into reasonably easy-to-run programs, which could come in handy in studies of the three-gluon vertex, and a number of other precision tests. No interface to fragmentation models exists so far. Program authors are recommended to give this some thought. It would be trivial to use independent fragmentation, while string or cluster fragmentation in principle would require additional colour information to be retained in the matrix element expressions. Approximate approaches could be found without too much ado, however.

There is little to be gained in going beyond third order matrix elements for LEP studies (at least, so long as loop calculations are lagging behind Born term ones), but general approximate expressions exist, and also a few exact ones, that could then be useful.

Parton Showers:
The modelling of parton showers is the central task in many QCD programs. Therefore

a wide selection of algorithms is available. There is also a continued improved understanding of the underlying perturbative QCD picture, which goes hand in hand with the improvement of the programs. It is probably true to say that four groups, more than others, have viewed their development work as long-term projects, and thus invested a significant amount of time to cover aspects not included in other programs. The four favoured programs are.

- HERWIG, which is the latest version in the Marchesini-Webber line of algorithms, and supersedes its (fairly similar) predecessors. The authors were the first to introduce angular ordering in Monte Carlos, and HERWIG still enjoys the status of the technically most advanced leading log shower program, also with many effects not part of standard leading log included. The detailed treatment of spin correlations is worth a special accolade.
- NLLjet, developed by Kato, Munehisa and coworkers, is the first full-scale program to offer parton shower evolution beyond the leading logarithmic approximation. This makes for a more exact and reliable description of the showering process.
- ARIADNE, developed by the Lund group, based on a picture originally proposed by the Leningrad group, is the only program not formulated in terms of parton branchings, but rather in terms of colour-anticolour dipole branchings. While the two pictures are formally equivalent to lowest order, the dipole scheme may come closer to a 'true' description of the underlying physics.
- JETSET, which is similar to HERWIG, but with important differences, and without many of the advanced features of HERWIG. However, it does contain some practical aspects, not found elsewhere, which come in handy in experimental studies.

In fact, these four algorithms, if combined with string fragmentation, tend to give an about equally good description of PEP and PETRA global event properties, according to studies in the literature. There is no reason to expect anything different at LEP, at least not until more detailed aspects of the programs are probed. This does not mean that all shower algorithms are equivalent; there certainly is a dividing line between coherent and conventional showering programs, but also coherent programs may have divergent predictions.

While it is clear that much interesting work exists, we still do not have one program that 'does it all'. The list of specific problems, either addressed only in a few programs, or not addressed at all, includes the following.

- Soft gluon interference effects. The angular ordering of subsequent branchings is here the most obvious manifestation, which is nowadays included in most programs. The related azimuthal asymmetries of soft gluon emission are only included in a few programs, however.
- Gluon polarization. Locally, this means that the production and decay planes of a gluon are related in a non-trivial manner. Also non-local correlations may arise; these are only covered by HERWIG.
- Lorentz covariance. The selection of a gauge for the evaluation of branching probabilities implies the need to pick a specific Lorentz frame. Therefore, no current showering algorithm is fully Lorentz covariant. The dipole algorithm of ARIADNE shows that it can be done, however.

- Hard emission. The leading log picture is based on a collinear approximation to the full kinematics. Therefore it is expected to describe the internal structure of jets well, but it is not guaranteed to give the proper rate of hard, non-collinear parton emission. As it happens, it is mainly the latter process which is probed by the standard subdivision of events into two-, three- or four-jet ones. Some programs (starting with JETSET) make use of the exact first order three-jet matrix element to adjust the probability for hard emission. This is mainly a patchwork solution of uncertain validity, but it is the best we know how to do today.
- Higher orders. Here, of course, NLLjet is the only program which has something to offer. While the Λ parameter of leading log programs stands in no simple relation to the fundamental $\Lambda_{\overline{MS}}$ one, the inclusion of loops in NLLjet implies a better defined Λ scale. Since the expansion is around a collinear emission picture, there is still no guarantee that hard, non-collinear emission is correctly reproduced.
- Shower cut-off. The shower evolution is cut off at some Q^2 scale, nowadays typically $Q_0^2 \simeq 1$ GeV^2, not because the branching probability becomes small, but because it becomes unmanageably large. It is assumed that emission below this scale can be included in the effective fragmentation picture, but we have some hints (e.g. the soft photon excess discussed in section 5.9) that this might not always be correct.

It is not obvious whether all the features above will be found in one single program in the near future, but, at least, the evolution of shower algorithms is certain to continue.

Parton Showers vs. Matrix Elements:
In order to make QCD generators more useful tools, it is probably necessary to combine the best features of the two approaches above, to obtain a consistent picture of both hard and soft parton emission. It is therefore recommended that Monte Carlo authors work towards a generator which, in a consistent manner, incorporates second order matrix elements, followed by next to leading order parton shower evolution. In addition, it would be desirable to have a complete $\mathcal{O}(\alpha_S^3)$ matrix element package.

Fragmentation:
The task of describing fragmentation is unquestionably the most difficult in QCD Monte Carlos, since the underlying QCD theory is not understood in the fragmentation region. Thus, contrary to most other topics, technical skills of Feynman graph calculation and the like are of no use; what is needed of program authors is a 'sound physical intuition', combined with a portion of luck.

Traditionally, one distinguishes three fragmentation schools.

- Independent fragmentation, here represented by COJETS, EURODEC and PARJET. Inside isolated jets, this picture is in reasonable agreement with data, but it is known to fail (for very fundamental reasons) in providing a complete description of event structures, in particular for the production of particles in between jets. Independent fragmentation should never be taken as the basic scenario, but can be useful to study variations.
- String fragmentation, represented by JETSET and programs based on JETSET: ARIADNE, NLLjet and TIPTOP. The development of the Lund string model, embodied in JETSET, probably represents the largest and most successful project of fragmentation studies. And yet, the final description is riddled with parameters, and leaves many questions unanswered.
- Cluster fragmentation is, in a fairly pure form, represented by HERWIG, and, in a hybrid string–cluster arrangement, by CALTECH-II. The cluster approach provides a viable alternative to strings, but some admixture of stringlike ideas seems to be required for successful phenomenology.

Although it is often possible to find good agreement between models and data, it is clear that we still do not have a fundamental understanding of the underlying physics. One may then worry about specific unsolved problems, like

- what is the actual mechanism of flavour production, in particular of baryon formation?
- how does intermittency arise?
- what is a correct description of Bose-Einstein effects?
- how shound one understand the soft photon and low mass dilepton excess observed in hadron collisions, and is it present also in e^+e^- annihilations?

and many more.

What we need, obviously, is to have QCD solved for us. Failing that, we need more alternative models, which can help us put things into perspective. One full-scale attempt, which does provide an alternative to conventional approaches, is the EPOS progam. DPSJET is an interesting toy model, and several ideas not yet implemented in complete programs could also be mentioned. However, the amount of current activity in the field of fragmentation studies compares rather unfavourably with the attention given e.g. to parton showers.

Decays:
The decaying of unstable hadrons into the observable final state particles is often considered trivial. Indeed, the construction of accurate and reliable decay routines is probably the most ungrateful undertaking in QCD Monte Carlo development. And yet, the task is neither simple nor unimportant.

Most programs do contain decay routines as a matter of course. Depending on which particular decay is considered, one may be better than another.

- Ordinary light hadron decays are based on known branching ratios, and so differences between programs should be minimal. EURODEC and JETSET may contain slightly more explicit matrix elements than do others.
- Charm and τ decays are particularly well modelled in EURODEC. This is (together with TIPTOP) the only program to keep track of polarization information, especially important for the τ.
- For bottom decays, experimental knowledge is limited, and the use of various phase-space models the only practical way out. Again EURODEC and JETSET should be about comparable.
- For top and other heavy objects, the decay into quarks is associated with non-negligible parton shower evolution. This is included e.g. in HERWIG and JETSET.

Standardization:
After the standard particle codes proposed by the Particle Data Group, we have here taken a second step by defining a set of commonblocks that should be used as a standard for all interfacing of event generators. It must be realized that the road does not end here, however, and that further standards will have to be defined in the future. As part of this evolution, Monte Carlo authors should strive to make their programs more modular.

Famous last words:
If, after having read this report, or even only the summary, you still do not know which one single program to use, then you may agree with our last conclusion:
Due to the large uncertainties present in any realistic QCD Monte Carlo, physics studies must be based on the use of at least two complete and independent programs.

Acknowledgements

All the authors of software mentioned in this report have been very helpful to us, by making pieces of text available, by carefully reading our reports, and/or by responding to questions of ours. We would in particular like to thank the following individuals: L. Angelini, F. Berends, S. Bethke, W. Giele, N. Glover, T. Kamae, K. Kato, J.H. Kühn, L. Lönnblad, N. Magnussen, G. Marchesini, C. Maxwell, D. Morris, T. Munehisa, R. Odorico, W. Ochs, J. Ranft, W.-K. Tung, G. Valenti, J. van der Bij, B.R. Webber, D. Zeppenfeld, and R.-y. Zhu.

In addition, we have profited from help, in various forms, from the following persons: P. Burrows, J. Layter, K. Moenig, J. Nicholls, G. Rudolph, R. Settles, and P. Zerwas.